Statistical analysis for the radiological characterization of radioactive waste in particle accelerators

Biagio Zaffora

le cnam

École Doctorale Sciences des Métiers de l'Ingénieur

Laboratoire Procédés et Ingénierie en Mécanique et Matériaux

THÈSE DE DOCTORAT

présentée par : **Biagio ZAFFORA**

soutenue le : **8 septembre 2017**

pour obtenir le grade de : **Docteur du Conservatoire National des Arts et Métiers**

Discipline : **Milieux Denses et Matériaux**

Spécialité : **Sciences des Matériaux**

Statistical analysis for the radiological characterization of radioactive waste in particle accelerators

THÈSE dirigée par

M. CHEVALIER Jean-Pierre	*Professeur du Cnam*
Mme. LUCCIONI Catherine	*Professeur des Universités, Cnam*

RAPPORTEURS

M. IOOSS Bertrand	*Chercheur Senior, EDF R&D*
M. LYOUSSI Abdallah	*Président du jury, Professeur, INSTN/CEA*

EXAMINATEURS

M. LYOUSSI Abdallah	*Président du jury, Professeur, INSTN/CEA*
M. DUTZER Michel	*Directeur Adjoint Développement et Innovation, ANDRA*
M. MAGISTRIS Matteo	*Co-encadrant, Physicien, CERN*
M. SAPORTA Gilbert	*Co-encadrant, Professeur Émérite, Cnam*

Abstract

This thesis introduces a new method to characterize metallic very-low-level radioactive waste produced at the European Organization for Nuclear Research (CERN). The method is based on: 1. the calculation of a preliminary radionuclide inventory, which is the list of the radionuclides that can be produced when particles interact with a surrounding medium, 2. the direct measurement of γ emitters and, 3. the quantification of pure-α, pure-β and low-energy X-ray emitters, called difficult-to-measure (DTM) radionuclides, using the so-called scaling factor (SF), correlation factor (CF) and mean activity (MA) techniques. The first stage of the characterization process is the calculation of the radionuclide inventory via either analytical or Monte Carlo codes. Once the preliminary radionuclide inventory is obtained, the γ-emitting radionuclides are measured via γ-ray spectrometry on each package of the waste population. The major γ-emitter, called key nuclide (KN), is also identified. The scaling factor method estimates the activity of DTM radionuclides by checking for a consistent and repeated relationship between the key nuclide and the activity of the difficult to measure radionuclides from samples. If a correlation exists the activity of DTM radiodionuclides can be evaluated using the scaling factor otherwise the mean activity from the samples collected is applied to the entire waste population. Finally, the correlation factor is used when the activity of pure-α, pure-β and low-energy X-ray emitters is so low that cannot be quantified using experimental values. In this case a theoretical correlation factor (CF) is obtained from the calculations to link the activity of the radionuclides we want to quantify and the activity of the key nuclide. The thesis describes in detail the characterization method, shows a complete case study and describes the industrial-scale application of the characterization method on over 1000 m^3 of radioactive waste, which was carried out at CERN between 2015 and 2017.

Keywords : Particle accelerator, radioactive waste, statistical analysis, scaling factor, activation

Résumé

Ce travail de thèse introduit une nouvelle méthode pour la caractérisation radiologique des déchets métalliques très faiblement radioactifs produits au sein de l'Organisation Européenne pour la Recherche Nucléaire (CERN). La méthode se base sur: 1. le calcul des radionucléides en présence, i.e. les radionucléides qui peuvent être produits lors de l'interaction des particules avec la matière et avec les structures environnantes les accélérateurs, 2. la mesure directe des émetteurs γ et, 3. la quantification des émetteurs α et β purs et de rayons X de faible énergie, appelés radionucléides difficile-a-mesurer (DTM), en utilisant les méthodes dites des "scaling factor" (SF), "correlation factor" (CF) et activité moyenne (MA). La première phase du processus de caractérisation est le calcul des radionucléides en présence à l'aide de codes de calcul analytiques ou Monte Carlo. Après le calcul de l'inventaire radiologique, les radionucléides émetteurs γ sont mesurés par spectrométrie γ dans chaque colis de la population. L'émetteur γ dominant, appelé "key nuclide" (KN) est identifié. La méthode dite des "scaling factors" permet d'estimer l'activité des radionucléides DTM après évaluation de la corrélation entre l'activité des DTM et l'activité de l'émetteur γ dominant obtenue à partir d'échantillons. Si une corrélation existe, l'activité des radionucléides DTM peut être évaluée grâce à des facteurs de corrélation expérimentaux appelés "scaling factors", sinon l'activité moyenne obtenue à partir d'échantillons prélevés dans la population est attribuée à chaque colis. Lorsque les activités des émetteurs α et β purs et des émetteurs X de faible énergie ne peuvent pas être estimées par mesure analytique la méthode des "correlation factors" s'applique. La méthode des "correlation factors" se base sur le calcul de corrélations théoriques entre l'émetteur γ dominant et les radionucléides de très faible activité. Cette thèse décrit en détail la nouvelle technique de caractérisation radiologique, montre un cas d'application complet et présente les résultats de l'industrialisation de la méthode ayant permis la caractérisation radiologique de plus de 1000 m^3 de déchets radioactifs au CERN entre 2015 et 2017.

Mots clés : Accélérateur de particules, déchets radioactifs, analyse statistique, scaling factor, activation

Résumé long en langue française

1. Contexte

Le CERN est un des plus grands laboratoires de physique du monde. Au sein du complexe des accélérateurs (Fig. 1), les particules sont accélérées à des énergies croissantes jusqu'à attendre 7 TeV dans le Grand Collisionneur de Hadrons.

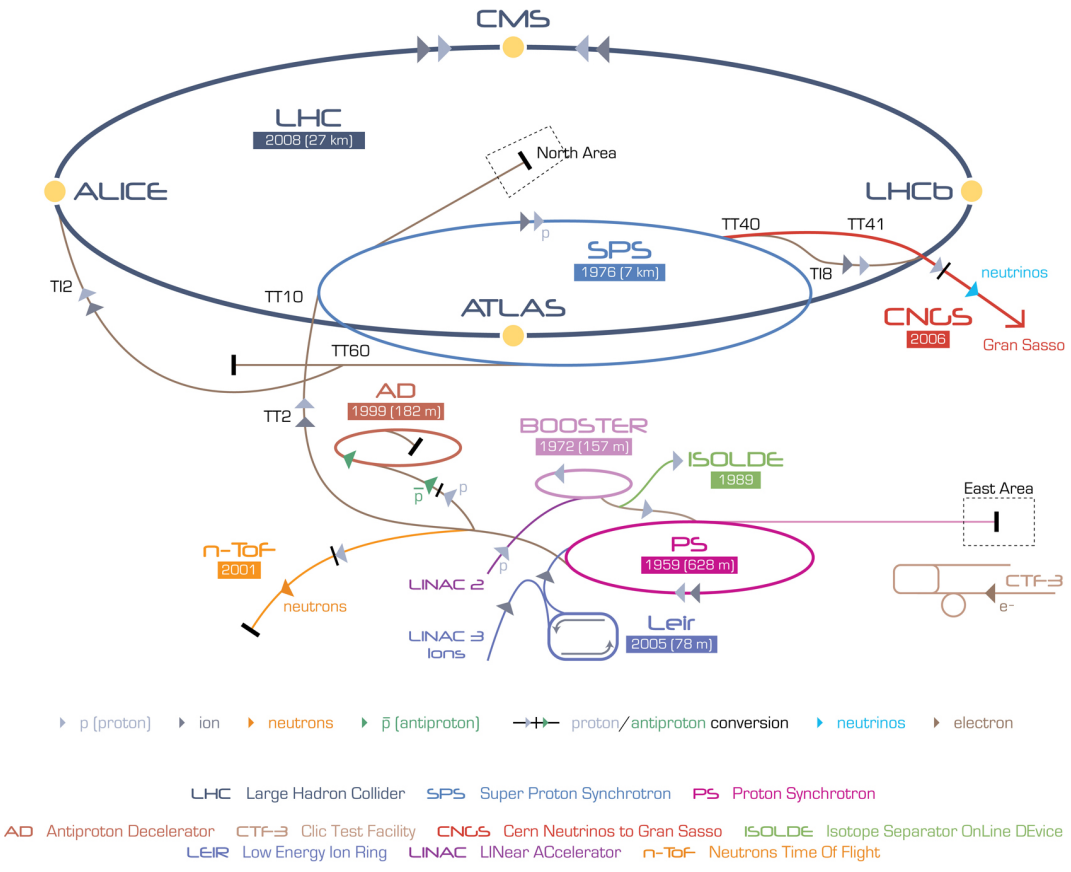


Figure 1: Complexe d'accélérateurs de particules du CERN.

Lorsque des particules interagissent avec les structures environnantes, un ou plusieurs mécanismes d'activation peuvent intervenir. En fin de vie, les matériaux activés peuvent être classés comme déchets radioactifs et doivent être caractérisés afin de démontrer leur acceptabilité au sein des centres de stockage.

Ce travail de thèse propose une nouvelle méthodologie de caractérisation radiologique des déchets de très faible activité (ou TFA) produits au CERN. Tout particulièrement, la méthode se concentre sur les déchets radioactif métalliques historiques dont peu ou pas d'informations sont disponibles au moment de leur caractérisation radiologique.

Le schéma global de la méthode de caractérisation est montré en Fig. 2. Les principales étapes du processus sont:

A. Identification d'une population de déchets radioactifs

B. Évaluation de l'inventaire radiologique via calculs analytiques ou simulations Monte Carlo

C. Quantification de l'activité des radionucléides émetteurs γ (ou ETM) par mesure directe

D. Test d'applicabilité des méthodes de ratios d'activité ou activité moyenne pour l'évaluation de l'activité des émetteurs β-purs

E. Quantification des facteurs de corrélation

F. Quantification de l'activité des radionucléides difficiles (DTM) ou impossibles (ITM) à mesurer

G. Estimation de l'incertitude

H. Évaluation de l'acceptabilité du lot de déchets radioactifs au sein du centre d'entreposage final par calcul de l'Indice Radiologique d'Acceptabilité en Stockage (IRAS).

L'Agence National Française pour la Gestion des Déchets Radioactifs (ANDRA) a établi des indices permettant de tester l'acceptabilité d'un colis et d'un lot de colis TFA en

entreposage final. Le premier indice, appelé IRAS, permets de tester l'acceptabilité d'un colis unique de déchets et est défini comme suit:

$$IRAS = \sum_i \frac{a_i}{10^{Classe_i}} \tag{1}$$

où a_i est l'activité massique du radionucléide i et $Classe_i$ renseigne sur la radio-toxicité de ce même radionucléide. Afin qu'un colis de déchets puisse être accepté en entreposage final il est nécessaire que l'IRAS soit inférieur à 10.

Ensuite l'ANDRA a aussi défini un indice fixant des limites d'acceptabilité sur un lot de colis de dechets radioactifs. L'IRAS colis est ainsi défini:

$$\langle IRAS \rangle = \frac{\sum_k M_k \cdot IRAS_k}{\sum_k M_k} \tag{2}$$

ou M_k est la masse du colis k et $IRAS_k$ est l'IRAS associé au colis k. Afin qu'un lot de déchets radioactifs soit acceptable en entreposage final l'IRAS lot doit être inférieur à 1.

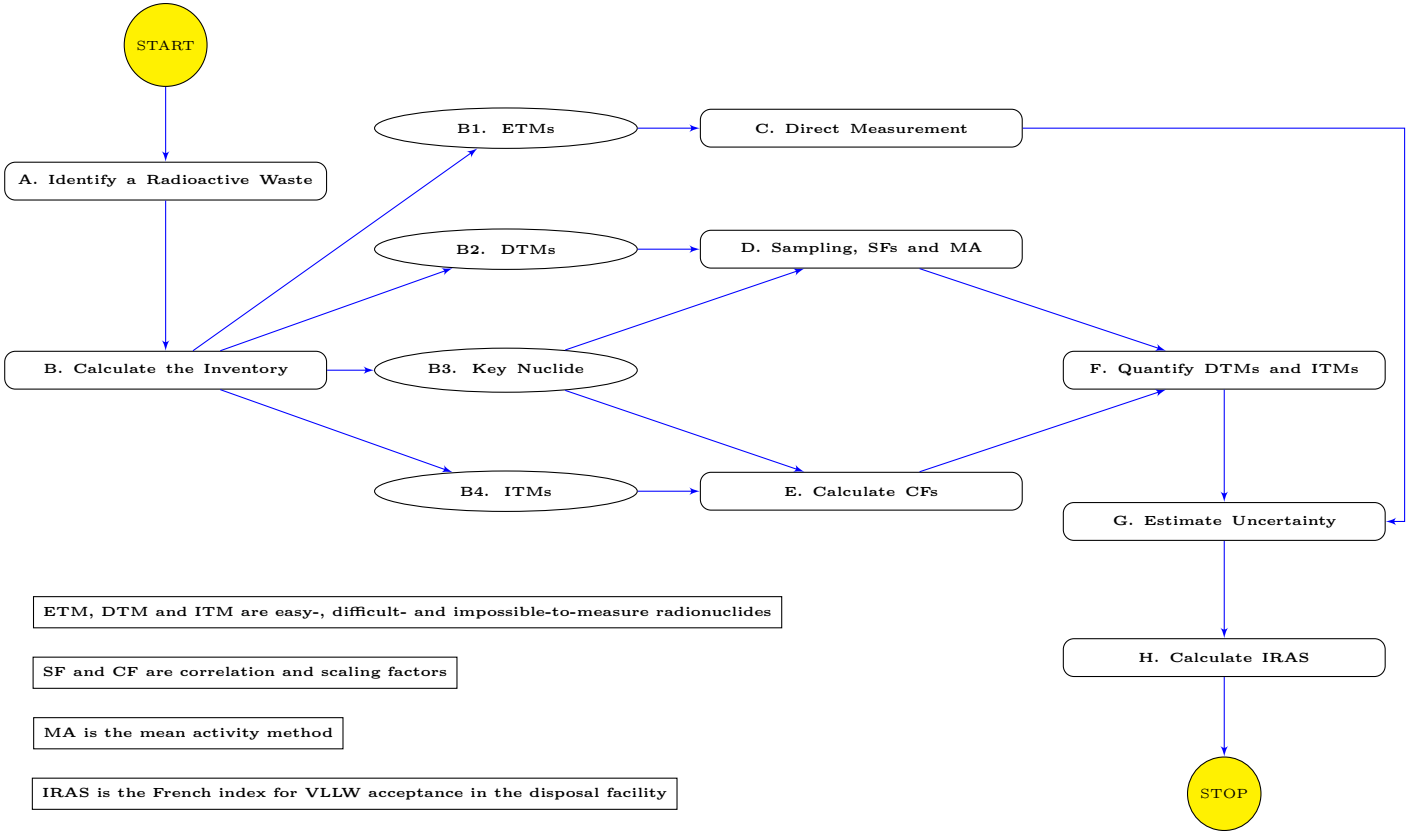


Figure 2: Processus de caractérisation radiologique des déchets métalliques TFA développé au CERN.

2. Évaluation de l'inventaire radiologique

L'inventaire radiologique est la liste des radionucléides produits avec une activité supérieure au seuil de déclaration (SD). Le seuil de déclaration est l'activité massique minimale à partir de laquelle un radionucléide doit être pris en compte lors du calcul de l'IRAS.

La méthodologie décrite dans cette thèse propose l'utilisation de codes de calcul afin d'établir l'inventaire radiologique. Tout particulièrement, afin de prendre en compte un nombre élevé de scenarios d'activation potentiels, un vecteur aléatoire **S** à cinq composants a été introduit:

$$\mathbf{S} = (\mathbf{CC}, E, P, t_i, t_c). \tag{3}$$

où:

- **CC** est un vecteur aléatoire représentant la composition chimique du matériau
- E est l'énergie de l'accélérateur
- P la position du matériau au sein du complexe d'accélérateurs
- t_i le temps d'irradiation et
- t_c le temps de décroissance.

Lorsque une unique combinaison des composants du vecteur aléatoire **S** est extraite, nous pouvons l'utiliser comme paramètre d'entrée d'un simulation de type Monte Carlo ou d'un calcul analytique.

Les simulations fournissent d'un coté la liste de tous les radionucléides qui peuvent être produits par activation et de l'autre des facteurs de corrélation permettant d'établir une relation fonctionnelle entre des radionucléides de type ETM (émetteurs γ) et des radionucléides DTM ou ITM.

La différence entre radionucléides DTM et ITM s'opère par le biais de leur contribution à l'IRAS C_i^{IRAS}, défini comme suit:

$$C_i^{IRAS} = \frac{\frac{a_i}{AL_i}}{\sum_j \frac{a_j}{AL_j}}. \tag{4}$$

avec AL étant égal à 10^{Classe}.

Tout particulièrement, si un radionucléide (hors ETM) contribue à l'IRAS plus de 1% alors il est defini difficile-à-mesurer (DTM). Lorsque la contribution d'un radionucléide est inférieure à 1% alors il sera classé comme impossible-à-mesurer (ITM).

La section suivante montre comment évaluer les activités des radionucléides DTM et ITM dans les déchets métalliques TFA produits au CERN.

3. Méthodes pour quantifier l'activité

Chaque radionucléide peut être classifié comme facile à mesurer (ETM), difficile (DTM) ou impossible à mesurer (ITM).

Un radionucléide est défini ETM s'il peut-être mesuré depuis l'extérieur d'un colis de déchets en utilisant des méthodes de mesure non-destructives, comme par exemple la spectrométrie γ.

Un radionucléide est défini DTM si la détermination de son activité massique nécessite des mesures de type destructive (dissolution acide, mesure par scintillation liquide ou comptage $\alpha\backslash\beta$).

Un radionucléide est défini ITM si son niveau d'activité massique est systématiquement inférieur aux seuils de détection des instruments communément employés pour leur mesure. L'activité massique de cette famille de radionucléides est généralement évaluée par simulation numérique.

Au CERN, la quantification de l'activité des ETM est faite grâce à des spectrométres γ portatifs, refroidis électriquement. Un exemplaire de ce détecteur est montré en Fig. 3.

Si un ETM a une demi-vie assez longue (par rapport à la durée du processus de

Figure 3: Le Falcon5k est un spectrométre γ portable, refroidi électriquement. Ce système de mesure est couramment employé au CERN pour quantifier l'activité des radionucléides ETM dans les colis de déchets radioactifs TFA.

caractérisation) et si son activité massique est corrélée à l'activité des DTM et ITM alors il est défini traceur (KN) et son activité peut etre utilisée pour quantifier l'activité massique des radionucléides DTM et ITM.

Une liste des radionucléides traceurs potentiels pour les déchets métalliques TFA produits au CERN est donnée dans la table qui suit.

Table 1: Liste de traceurs potentiel pour les déchets métalliques TFA. En gras sont identifiés les KN utilisés au CERN à la date d'écriture de cette thèse.

Matériau	**KNs**	**$T_{1/2}$ (ans)**	**Émetteur γ principal**
Acier	Na-22	2.603	1275 keV
	Ti-44	58.9	1157 keV (Sc-44)
	Co-60	5.2711	1173 keV, 1332 keV
Cuivre	Na-22	2.603	1275 keV
	Ti-44	58.9	1157 keV (Sc-44)
	Co-60	5.2711	1173 keV, 1332 keV
	Rh-101	3.3	127 keV, 198 keV, 325 keV
	Sb-125	2.7586	428 keV, 601 keV, 636 keV
Aluminium	**Na-22**	2.603	1275 keV
	Ti-44	58.9	1157 keV (Sc-44)
	Co-60	5.2711	1173 keV, 1332 keV

L'activité massique des radionucléides DTM peut-être calculée grâce aux méthodes

des ratios d'activité ou de l'activité moyenne. Ces méthodes se basent sur le prélèvement d'échantillons représentatifs à partir de la population de déchets. Chaque échantillon est mesuré par spectrométrie γ et par radiochimie et le ratio d'activité (SF) est calculé:

$$SF_i = \frac{a_{DTM_i}}{a_{KN_i}}. \tag{5}$$

où a_{DTM_i} est l'activité massique du DTM dans l'échantillon i et a_{KN_i} est l'activité massique du traceur dans le même échantillon.

Une distribution des SF_i peut ensuite être construite. Si les activités des DTM et KN sont corrélées (en général le coefficient de corrélation doit être supérieur à 0.5) alors il est possible choisir une estimateur de tendance centrale approprié selon la distribution expérimentale obtenue.

La pratique expérimentale montre que, souvent, la moyenne géométrique est un estimateur robuste pour les ratios d'activité. La moyenne géométrique des ratios d'activité $\overline{G}_{SF}$ peut être calculée à l'aide de la formule suivante:

$$\overline{G}_{SF} = exp\left(\frac{1}{n}\sum_{i=1}^{n}\ln(SF_i)\right) = \left(\prod_{i=1}^{n} SF_i\right)^{\frac{1}{n}}. \tag{6}$$

Si nous indiquons avec SF le ratio d'activité obtenu à partir des échantillons, alors l'activité massique du DTM i dans le colis j ($\hat{a}_{DTM_{i,j}}$) peut-être calculée par le biais de la formule suivante:

$$\hat{a}_{DTM_{i,j}} = SF \times a_{KN_j} \tag{7}$$

où a_{KN_j} est l'activité massique du KN mesurée dans le colis j.

Lorsque la corrélation entre les activités des DTM et KN est inférieure à 0.5, la méthode de la moyenne s'applique. Cette méthode consiste a attribuer à chaque colis l'activité moyenne obtenue à partir de la mesure des échantillons.

Pour conclure, l'activité massique des radionucléides ITM est estimée en applicant un processus similaire de celui décrit pour les radionucléides DTM. La différence réside dans

l'estimation des ratios d'activité que, dans le cas des ITM, est obtenue à partir de codes de calcul.

Si CF indique le ratio d'activité (estimé par calcul) corrélant l'activité des ITM avec le KN, alors l'activité du radionucléide ITM dans le colis j ($\hat{a}_{ITM_j}$) est donnée par:

$$\hat{a}_{ITM_j} = CF \times a_{KN_j}. \tag{8}$$

où a_{KN_j} est l'activité massique du KN mesurée dans le colis j.

4. Estimation des incertitudes

Afin d'assurer le respect des critères d'acceptation fournis par l'ANDRA il est essentiel estimer les incertitudes associées à l'IRAS (colis et lot).

Ces incertitudes peuvent être évaluées en utilisant les formules classiques de propagation des incertitudes.

L'incertitude combinée de l'IRAS d'un colis est donnée par:

$$u_c^2(IRAS) = \sum_{i=1}^{n} \frac{u^2(a_i)}{AL_i^2} + \sum_{\substack{i=1 \\ i\neq j}}^{n} \sum_{\substack{j=1 \\ j\neq i}}^{n} \frac{r(a_i, a_j) \cdot u(a_i) \cdot u(a_j)}{AL_i \cdot AL_j} \tag{9}$$

où $u(a_i)$ et $u(a_j)$ sont les incertitudes associées aux activités des radionucléides i et j, $r(a_i, a_j)$ est la corrélation entre leur activités et AL_i et AL_j sont les limites d'activité des radionucléides i et j ($AL = 10^{Classe}$).

L'incertitude de l'IRAS d'un lot s'obtient par propagation des incertitudes pour des variables non corrélées. Cette hypothèse est valable car l'IRAS d'un colis et le poids sont des quantités non-corrélées. Tout particulièrement, nous obtenons:

$$u_c^2(\langle IRAS \rangle) \simeq \sum_{k=1}^{n} \left(\frac{M_k}{M_{tot}} \right)^2 \cdot u^2(IRAS_k) + \sum_{k=1}^{n} \left(\frac{IRAS_k}{M_{tot}} \right)^2 \cdot u^2(M_k) \tag{10}$$

où M_{tot} est la masse totale du lot de colis et $u^2(IRAS_k)$ et $u^2(M_k)$ sont les incertitudes au carré de l'IRAS et du poids du k^{ième} colis.

Si l'on divise Eq. 10 par $\langle IRAS\rangle^2$ nous pouvons reformuler l'incertitude comme suit:

$$\left(\frac{u(\langle IRAS\rangle)}{\langle IRAS\rangle}\right)^2 \simeq \sum_{k=1}^{n}\left(\frac{u(IRAS_k)}{IRAS_k}\right)^2 + \sum_{k=1}^{n}\left(\frac{u(M_k)}{M_k}\right)^2. \tag{11}$$

Si les incertitudes des $IRAS$ colis sont similaires ($u(IRAS_k) = u(IRAS)$) ainsi que les incertitudes des poids des colis ($u(M_k) = u(M)$) l'equation de l'incertitude de l'IRAS lot peut etre simplifiée.

En utilisant la definition de variance nous pouvons écrire:

$$\sum_{k=1}^{n} M_k^2 = n\overline{M}^2 + nVar(M) \simeq n\overline{M}^2 + ns^2(M) \tag{12}$$

et

$$\sum_{k=1}^{n} IRAS_k^2 = n\overline{IRAS}^2 + nVar(IRAS) \simeq n\overline{IRAS}^2 + ns^2(IRAS) \tag{13}$$

où s^2 represente le carré de l'écart-type expérimentale de M et $IRAS$, et $\overline{M}$ et $\overline{IRAS}$ sont les poids et $IRAS$ moyens.

Si nous injectons les deux dernières équations dans la formule générale de l'incertitude de l'IRAS d'un lot de déchets nous pouvons calculer:

$$\begin{aligned} u_c^2(\langle IRAS\rangle) \simeq & \frac{1}{M_{tot}^2}\left[u^2(IRAS)(n\overline{M}^2 + ns^2(M))\right] + \\ & \frac{1}{M_{tot}^2}\left[u^2(M)(n\overline{IRAS}^2 + ns^2(IRAS))\right]. \end{aligned} \tag{14}$$

En fin, si on écrit $M_{tot}^2 = n\overline{M}^2$ nous obtenons la formule définitive pour calculer l'incertitude associée à l'IRAS d'un lot de déchets radioactifs:

$$\begin{aligned} u_c^2(\langle IRAS\rangle) \simeq & \frac{u^2(IRAS)}{n}\left(1 + \frac{s^2(M)}{\overline{M}^2}\right) + \\ & \frac{u^2(M)}{\overline{M}^2}\left(\frac{\overline{IRAS}^2}{n} + \frac{s^2(IRAS)}{n}\right). \end{aligned} \tag{15}$$

5. Résultats

La méthode de caractérisation proposée dans cette thèse a été testée et validée dans le cadre du projet SHERPA, le plus grand projet d'élimination de déchets radioactifs TFA réalisé au CERN entre 2014 et 2017.

L'objectif du projet SHERPA est celui de créer un processus de caractérisation, traitement et élimination de plus de 1000m^3 de déchets radioactifs métalliques TFA. Le sommaire du nombre de colis caractérisés dans le cadre du projet SHERPA est présenté en Tab. 2.

Table 2: Sommaire des colis traités dans le cadre du projet SHERPA jusqu'à juin 2017. 80% de ces colis sont en entreposage définitif. La partie restante est caractérisée et prête pour l'expédition.

Matériau	**N. colis**	**Poids (en tonnes)**
Aluminium	68	81.6
Cuivre	17	38.7
Acier	262	572.4
Total	347	692.7

Un exemple de colis en cours de préparation pour l'expédition est donné en Fig. 4.

Figure 4: Chargement d'un lot de colis de déchets SHERPA pour expédition en entreposage.

Dans le cadre du projet SHERPA l'inventaire radiologique a été obtenu en considérant

un nombre élevé de scénarios d'activation (>2.35 millions), pour tous les matériaux couramment utilisés au CERN, toutes les énergies des accélérateurs et pour des temps d'irradiation et de décroissance allant de 1 jusqu'à 40 ans.

L'inventaire radiologique obtenu pour les familles de métaux plus courants est donné en Tab. 3.

Table 3: Inventaire radiologique pour les familles de métaux traités dans le cadre du projet SHERPA. Environ 2.35 millions de scénarios d'activation ont été pris en compte.

Famille	**ETM**	**DTM/ITM**
Aluminium	Na-22, Al-26, Ti-44, Mn-54 Co-57, Co-60, Zn-65	H-3, C-14, Cl-36, Ar-39 Fe-55, Ni-63
Cuivre	Na-22, Ti-44, Mn-54, Co-57 Co-60, Zn-65, Mo-93, Rh-101 Ag-108m, Ag-110m, Sn-121m, Sb-125 Gd-148, Hg-194, Bi-207	H-3, C-14, Cl-36, Ar-39 Ca-41, Fe-55, Ni-63
Acier	Na-22, Ti-44, Mn-54, Co-57 Co-60, Zn-65, Mo-93, Nb-93m Nb-94, Tc-99	H-3, Be-10, C-14, Cl-36 Ar-39, Ca-41, Fe-55 Ni-63

De tous les radionucléides prévus par calcul, seulement une partie dépasse les seuils de déclaration et doit donc être utilisée pour le calcul de l'IRAS. Les ETM dépassant les seuils de déclaration sont indiqués dans en Tab. 4. Les distributions d'activité massique pour Co-60 et Na-22 sont montrées en Fig. 5.

Table 4: Liste des radionucléides ETM dépassant les seuils de déclaration dans le cadre du projet SHERPA.

Matériau	**ETM**
Aluminium	Na-22, Ti-44, Co-60, Ag-108m
Cuivre	Ti-44, Co-60, Ag-108m
Acier	Ti-44, Mn-54, Co-60, Zn-65, Ag-108m, Bi-207

Les radionucléides DTM dépassant les seuils de déclaration sont H-3, Fe-55 et Ni-63. La distribution des activités massiques du H-3 dans l'aluminium est montrée comme un exemple en Fig. 6.

Pour conclure, les ITM détectés au delà du seuil de déclaration sont C-14 (19 colis), Cl-36 (19 colis) et Ar-39 (15 colis). Cependant, leur impact sur l'IRAS est négligeable.

Après avoir calculé l'inventaire radiologique et avoir estimé les activités des différents

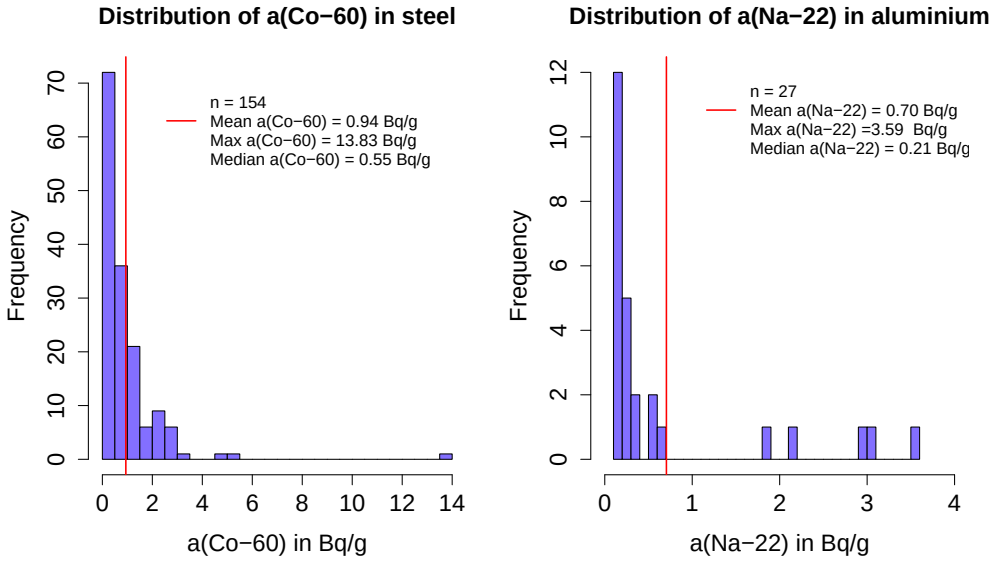


Figure 5: Histogrammes des activités massiques pour Co-60 et Na-22 en acier et aluminium. La figure montre seulement les activités qui dépassent les seuils de déclaration. Les lignes rouges indiquent les activités moyennes.

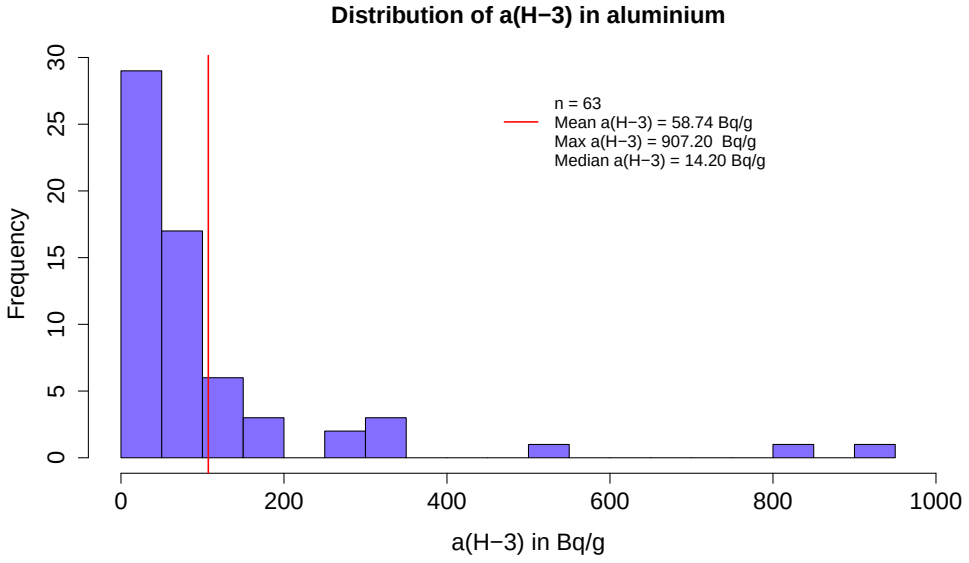


Figure 6: Histogramme de l'activité spécifique du H-3 dans des déchets d'aluminium. Seulement les valeurs dépassant les seuils de déclaration sont montrées. La ligne rouge indique l'activité massique moyenne du H-3.

radionucléides en présence, il est possible d'évaluer l'IRAS des colis et l'IRAS lot. Les graphiques qui suivent présentent les sommaires des IRAS colis pour le projet SHERPA

et par famille de matériau. La contribution de chaque radionucléide à l'IRAS est aussi donnée.

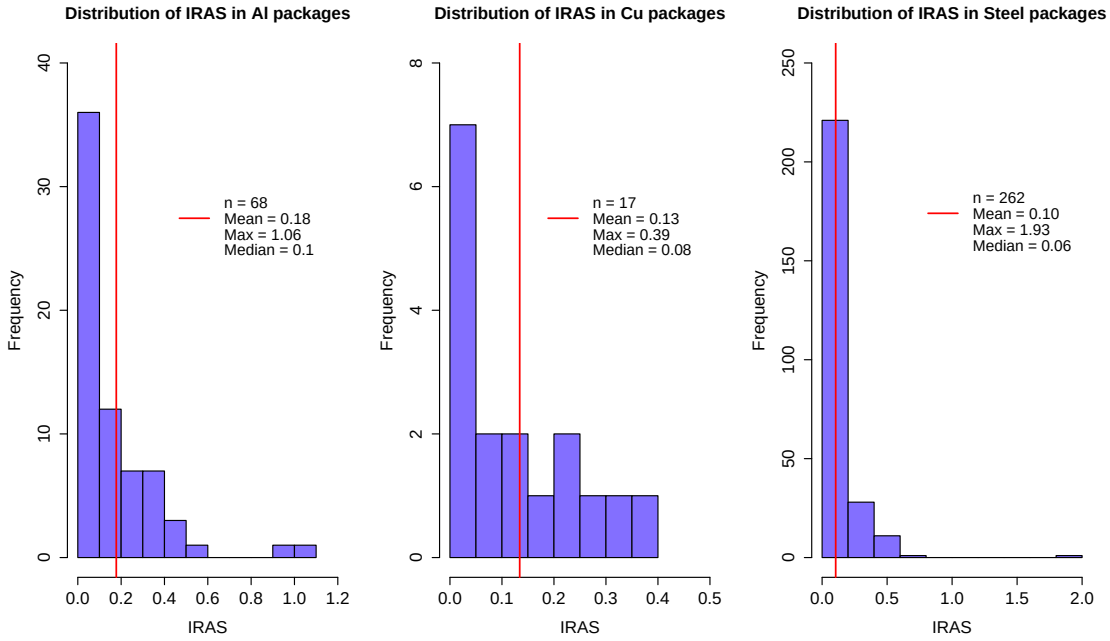


Figure 7: Histogrammes des IRAS colis pour les trois familles de metaux du projet SHERPA. Les lignes rouges indiquent les valeurs moyenne de l'IRAS colis.

6. Conclusions

L'objectif de ce travail de thèse est celui de proposer une solution pour la caractérisation radiologique des déchets métalliques TFA produits au sein des accélérateurs de particules. La méthode proposée se base sur une combinaison de simulations numériques et mesures permettant d'établir l'inventaire radiologique, d'évaluer l'activité des radionucléides plus importants et d'évaluer l'acceptabilité des colis de déchets dans les centres d'entreposage à partir du calcul de l'indice IRAS.

Le processus de caractérisation ici décrit est aujourd'hui employé en routine pour la caractérisation des déchets radioactifs métalliques produits au CERN. Cette méthode est acceptée par l'ANDRA, est applicable à la majorité de déchets métalliques et peut être facilement adaptée à d'autre familles de déchets.

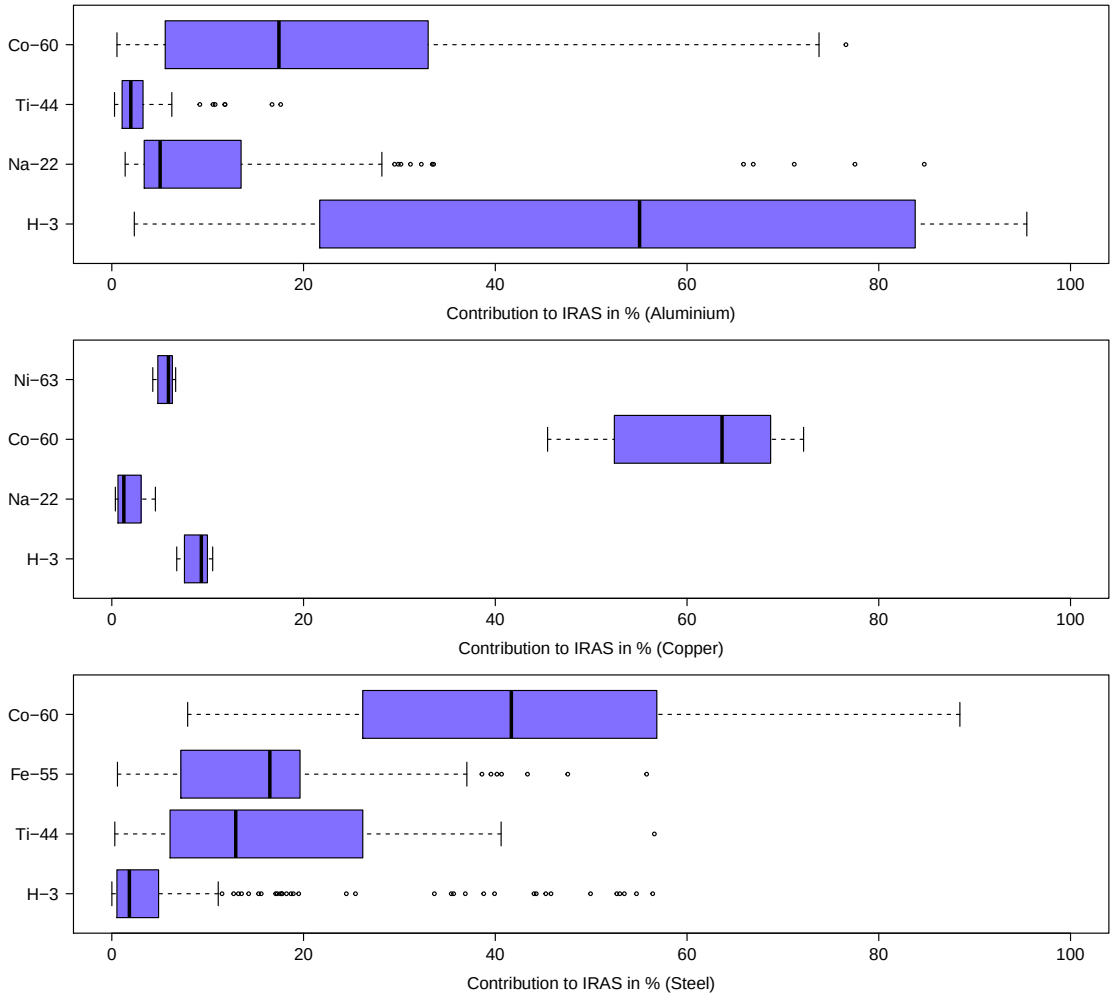


Figure 8: Contribution des radionucléides principaux à l'IRAS pour les 347 colis de déchets radioactifs du projet SHERPA traités avant juin 2017.

Le travail de cette thèse a aussi permis d'augmenter la taille des déchets caractérisés et éliminés au CERN: nous sommes passés de la caractérisation de peu d'objets avec une histoire radiologique connues à des centaines de tonnes de déchets avec histoire radiologique peu ou pas connue.

Nous avons aussi démontré que l'impact des éléments trace dans les matériaux est négligeable en termes d'inventaire radiologique.

La partie centrale du travail de thèse regarde l'adaptation de la méthode des ratios d'activité aux besoins du CERN. Cette méthode a en effet été adaptée et validée et est aujourd'hui à la base de la quantification de l'activité des radionucléides DTM.

Un point de force de la nouvelle méthode de caractérisation réside dans sa capacité de concentrer l'effort de caractérisation sur les radionucléides qui contribuent le plus à l'IRAS. Pour ce faire, des critères statistiques ont été développés et utilisés. Cela a en effet permis d'appliquer les méthodes de caractérisation plus longues et couteuse seulement aux radionucléides qui ont un impact avéré sur l'IRAS.

Pour conclure, la méthode de caractérisation proposée dans cette thèse est aujourd'hui implémentée à une échelle industrielle au CERN. La méthode a permis la caractérisation ou élimination d'environ 690 tonnes de déchets métalliques TFA entre 2015 et 2017 et peut être facilement adaptée à la caractérisation de nouvelles familles de déchets radioactifs.

Acknowledgements

I would like to express my personal thanks to my CERN supervisor Dr. Matteo Magistris because his guidance has been warm and illuminating. He is an invaluable mentor, kind and has always kept a sense of humour when I had lost mine. Today, thanks to his continuous example, I am a better person.

My thanks go also to Dr. Luisa Ulrici, the head of CERN radioactive waste section. This work would have not seen the light of day without her support and sponsorship.

I am infinitely grateful to my professors Catherine Luccioni, Jean-Pierre Chevalier and Gilbert Saporta. Their advice has always been precise, concrete and, most of all, useful.

Pr. Luccioni has followed closely my career since the very beginning, when I was a young γ-spectrometry technician. Today, after many years, her wise presence is still comforting and encouraging. In addition, her attention to details drove me to learn to finally punctuate my prose better.

Pr. Jean-Pierre Chevalier has expertly guided me through my PhD journey. His advice to "focus on what is really important" is discreetly pinned over my desk and has been a guidance for my work and my life until this day.

I am grateful, beyond words, to Pr. Gilbert Saporta for helping me to discover the beautiful world of statistics and data analysis. If today I am pondering about a career in statistics it is surely because of his example.

I would like to express my deepest gratitude to Pr. Abdallah Lyoussi, Dr. Bertrand Iooss and Mr. Michel Dutzer for accepting to be part of the committee of my PhD thesis. Their comments were very precious and have allowed me to look at my work from a different perspective.

I would like to finish by thanking my family, friends and colleagues. They created the perfect environment for me to reach this goal.

First of all my wife Julia, for her constant support and for reminding me that work is only one aspect of life and that I am surrounded by a lot of beautiful things to look at.

Stealing the words of a far better person than myself, I would like to thank my parents for the greatest gift: poverty. If that is the baseline, everything else is an enjoyable conquest to watch with the eyes of wonder.

A special thanks goes to my colleagues and friends Elpida, Francesco and Sven. They made my lunches less lonely, my weekends funnier and my life richer by far.

To conclude, thanks to Asimov, Bufalino, Dostoevsky, Eco, Gary, Kristof, Marquez, Murakami, Pamuk, Rand, Rushdie, Saramago, Singer and the many many others whose literature has allowed me to enjoy countless, sleepless nights and to visit beautiful, unknown worlds.

Contents

List of Tables

List of Figures

Acronyms

AA	Antiproton Accumulator
AC	Antiproton Collector
AD	Antiproton Decelerator
ALICE	A Large Ion Collider Experiment
ANDRA	French National Agency for Radioactive Waste Management
ATLAS	A Toroidal LHC ApparatuS
CERN	European Organization for Nuclear Research
CF	Correlation factor
CMS	Compact Muon Solenoid
CNGS	CERN Neutrinos to Gran Sasso
COMPASS	COmmon Muon and Proton Apparatus for Structure and Spectroscopy
CS	Surface Contamination
DT	Declaration Threshold
DTM	Difficult-to-measure
ETM	Easy-to-measure
FOPH	Federal Office of Public Health
GUM	Guide to the expression of uncertainty in measurement
IAEA	International Atomic Energy Agency
IRAS	Indice Radiologique d'Acceptabilité en Stockage
ISOLDE	Isotope mass Separator On-Line facility
ISR	Intersecting Storage Ring
ITM	Impossible-to-measure
JCGM	Joint Committee for Guides in Metrology
KN	Key nuclide
LE	Exemption Limit
LEAR	Low Energy Antiproton Ring
LEIR	Low Energy Ion Ring
LEP	Large Electron-Positron collider
LSC	Liquid scintillation counting
LHC	Large Hadron Collider
LHCb	Large Hadron Collider beauty
MA	Mean activity method
MCA	Multi-channel analyser
MDA	Minimum Detectable Activity
nTOF	neutron Time-of-Flight

ORaP	Ordonnance sur la Radioprotection
PS	Proton Synchrotron
SC	Synchro-Cyclotron
SF	Scaling factor
SPS	Super Proton Synchrotron
TFA	Trés Faible Activité
VLLW	Very-low-level waste
WANF	West Area Neutrino Facility

Introduction

If knowledge can create problems, it is not through ignorance that we can solve them.

Isaac Asimov
Asimov's guide to science.

Radioactive waste is produced at particle accelerators as a consequence of the interaction of particle beams with matter. The principal mechanism of waste production in these facilities is the activation of the materials surrounding the particle beams. Neutrons, protons, pions and photons can interact with matter and produce unstable nuclei which will lose their energy in excess via radioactive decay.

Activated materials that cannot be reused or recycled need to be disposed of in dedicated disposal facilities. One of the steps needed to dispose of radioactive waste is the so-called radiological characterization. The radiological characterization of a waste consists of establishing the list of produced radionuclides and evaluating their activity.

In the present thesis we introduce a new characterization strategy, based on numerical experiments and measurements, that allows us to safely dispose of legacy very-low-level waste (VLLW) produced at CERN. According to the definition of the International Atomic Energy Agency (IAEA), VLLW is "waste that does not need a high level of containment and isolation and, therefore, is suitable for disposal in near surface landfill type facilities with limited regulatory control" [IAEA 2009a]. VLLW represents the vast majority of the total amount of radioactive waste currently stored at CERN and that will be produced in the future. A specific methodology to characterize this family of waste was therefore needed.

The process we developed lies on extensive Monte Carlo and analytical calculations

performed to estimate reference radiological inventories for common materials used at particle accelerators such as copper, steel and aluminium. We use destructive and non-destructive nuclear measurements to quantify the activities of the radionuclides and to test for their correlation. Some radionuclides are easily measurable using non-destructive assay methods and we will refer to them as ETM radionuclides. In the context of this work ETM radionuclides are the ones measurable via γ-ray spectrometry. Common examples of ETMs encountered on radioactive metallic waste are Co-60, Na-22 and Ti-44. Radionuclides that require complex destructive methods for the quantification of their activity are called difficult-to-measure (DTM). This second category mainly includes pure β-emitters, such as Ni-63, and low-energy X-ray emitters, such as Fe-55.

Various techniques were developed for the radiological characterization of waste produced in nuclear power plants. One can cite the known scaling factor (SF) and correlation factor (CF) methods. These methods are standardized and useful information can be found in a number of references such as [IAEA 2007] [ISO 2007] [ISO 2013] [IAEA 2009b]. We adapted and implemented some of these methods to take into account the peculiarities of radioactive waste from particle accelerators.

The main result of the present work is the design and implementation of a procedure to radiologically characterize very-low-level waste produced at CERN. The methodology is accepted by the French National Agency for Radioactive Waste Management (ANDRA) and has allowed the characterization of more than 690 tons of activated metals between 2015-2017.

This thesis consists of 6 chapters. Chapter 1 describes the context of the present study. It introduces the CERN's accelerator complex and the beam losses at the origin of the activation mechanisms. After the introduction of the terminology used in the thesis, the chapter discusses the technical requirements that a waste must respect to be accepted into a final disposal facility. A list of the major waste populations that are presently stored at CERN is then given. The chapter ends with a general overview of the radiological characterization process we developed.

Chapter 2 deals with the calculations performed to estimate the radionuclide inventory of activated materials. After that the perimeter of a waste population is defined, information

about its origin, its chemical composition, its volume and its radiological history is collected. This data is crucial to estimate preliminary radiological inventories. For legacy waste such a set of information is often missing and, only in the best cases, partial or incomplete. In this chapter we introduce a numerical method that allows us to make predictions even if the input data is limited or missing.

Chapter 3 describes the instrumental and analytical techniques used to quantify the activity of the radionuclides in a waste population. We introduce γ-ray spectrometry for the measurement of ETM radionuclides and the scaling factor (SF) and the mean activity (MA) methods to estimate the activity of DTM radionuclides. Some difficult-to-measure radionuclides have such a low activity that often cannot be measured within a reasonable time/cost framework. Moreover, their activity is often below the detection capability of the common instruments used for their quantification. The IAEA decided to name this category of radionuclides impossible-to-measure (ITM) [IAEA 2007].

From a practical point of view ITM radionuclides are similar to DTMs because they cannot be measured from outside a waste package using non-destructive assay techniques. ITMs however differ from DTMs because their activity can only be estimated via analytical calculations. In Ch. 3 we also discuss the difference between DTMs and ITMs, we introduce some techniques to sample radioactive waste and we shortly describe how statistical learning approaches can be used to estimate theoretical correlation factors between ETM and DTM radionuclides.

Chapter 4 deals with the quantification of uncertainty in the characterization process. We list the major terms that contribute to the uncertainty budget following the recommendations of the Guide to the expression of uncertainty in measurement (GUM) [JCGM 2008a]. For some terms of the uncertainty budget we cannot apply the classical techniques used to quantify uncertainties. For these cases we implemented a Monte Carlo method and applied statistical learning techniques that can be used in support of the uncertainty evaluation.

Chapter 5 is divided into three major sections. The first part describes the application of the radiological characterization procedure to a population of activated copper. The waste consists of $\sim$8.5 tons of copper recovered after shredding historical cables extracted

from the accelerators tunnels. We show how to establish the radionuclide inventory, how to calculate scaling and correlation factors and how the specific activity of each radionuclide is estimated using both measurements and calculations. We finally present the calculation of the parameters needed to evaluate the acceptability of the waste in the disposal facility.

The second section of the chapter describes the operational results obtained applying the new characterization method at CERN, with a focus on the metallic waste disposed of in 2015-2017. We present a feedback and the lessons learned from the disposal campaigns completed during this period.

We end the chapter with an overview of key operational parameters which are of interests for the characterization of waste. We finally list and discuss identified limitations of the methodology, ongoing improvements and axes of future research.

The last chapter of the thesis presents a summary of the major stages of the characterization process, the findings of the work performed and some critical points that we believe should be carefully analysed when characterizing radioactive waste. We conclude the thesis outlining the ongoing and future applications of the radiological characterization method at CERN.

Chapter 1

General context

I am quite surprised that it happened during my lifetime. It is nice to be right about something sometimes.

Peter Higgs
Comment after the discovery of the Higgs boson.

Introduction

CERN was founded in 1952 with the mandate of establishing a fundamental physics laboratory in the Europe emerged from the Second World War. Since then, the objective of fundamental physics has evolved at levels of subnuclear particles. Today the main area of research of the laboratory is the study of the fundamental constituents of matter and the forces acting between them. Because of this evolution, CERN is also referred to as the European Laboratory for Particle Physics [CERN 2015].

Over the last 60 years, particle accelerators of increasing energy have been thought, designed, built and put into operation at CERN. However, due to the interaction of particle beams with matter, part of the accelerator structure and its surrounding has become radioactive.

The complex mechanism of the activation process depends on the characteristics of the radiation environment, which is determined by the energy and the type of accelerated particles, such as electrons, protons and heavy ions.

This chapter begins with a description of the most important machines at CERN, with a particular focus on present and past accelerators. A preliminary understanding of machine characteristics is useful to explain the common activation mechanisms leading to the production of radioactive waste.

The second section focuses on the mechanism of material activation. Activation occurs when particles interact inelastically with matter. At CERN, the particles responsible for activation are the primary particles (beam losses) and the secondary particles they generate when interacting with the accelerator components. The loss mechanisms are explained here together with the typical particle spectra encountered in CERN's accelerators. The general formula of activation is also given.

In the third section, the radioactive waste disposal pathways presently available at CERN are described. CERN is located on the border across two host states (France and Switzerland) and a so-called "Tripartite Agreement" [Acc 2011] defines the framework for the elimination of radioactive waste towards host states. An overview of the principal guidelines for the elimination of radioactive waste is given. The last part of the section shortly lists the inventory of radioactive waste stored at CERN at the date of writing (June 2017).

The fourth section of the present chapter focuses on the main topic of this study: the radiological characterization of radioactive waste. In this section we give an overview of the characterization workflow, we introduce the essential terminology and describe the structure of the thesis.

1.1 CERN's accelerator complex

The accelerator complex at CERN is a succession of machines that accelerate particles to increasingly higher energies as showed in Fig. 1.1[1].

At the entrance of the acceleration chain, protons are generated in a source consisting of hydrogen gas. Each machine enhances the energy of the beam of particles before the injection into the next accelerator of the facility. The last element of the chain is the LHC

[1] ©2013-2015 CERN.
`http://te-epc-lpc.web.cern.ch/te-epc-lpc/machines/lhc/general.stm`, 7 October 2015.

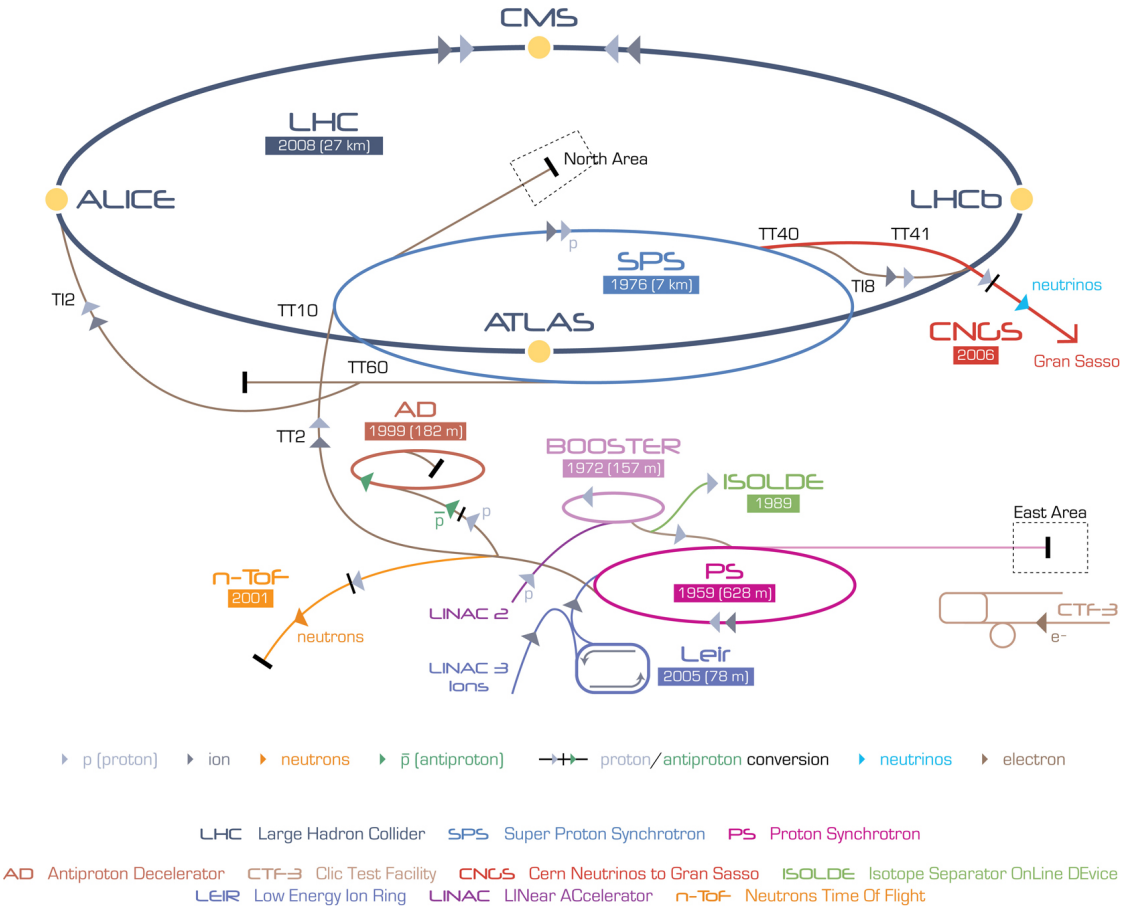


Figure 1.1: Schematic layout of the CERN's accelerator complex.

in which particle beams can be accelerated to the energy of 7 TeV per beam [CERN 2015].

Linac 2, the first accelerator in the chain, accelerates the protons up to the energy of 50 MeV. This machine started up in 1978, when it took the place of Linac 1, and will be definitely replaced by Linac 4 between 2017 and 2018. A section of the Linac 2 is showed in Fig. 1.2[2]

Figure 1.2: Linac 2, at the beginning of the CERN's accelerator chain.

In Linac 2 protons pass through conductors, alternately charged positive or negative, causing their acceleration. Quadrupole magnets ensure the protons to remain focused in the beam axis. At the end of the machine the protons gain 5% in relativistic mass. Finally the accelerated particles enter the Proton Synchrotron (PS) Booster, the next step in CERN's accelerator chain.

The PS Booster, accelerates protons up to 1.4 GeV. This machine is made of four superimposed synchrotron rings of 25 m of diameter. The beam leaving the PS Booster is injected either into the Proton Synchrotron (PS) or directly into ISOLDE (Isotope mass Separator On-Line facility) for experiments in nuclear physics, such as for example the study of exotic nuclei, far from the stability valley [CERN 2015]. A section of the PS

[2]©2013-2015 CERN.
`http://home.web.cern.ch/about/accelerators/linear-accelerator-2`, 7 October 2015.

Booster is showed in Fig. 1.3[3].

Figure 1.3: A section of the PS Booster.

The next accelerating stage is the PS, which pushes the beam up to 25 GeV per proton. The PS accelerates either protons delivered by the PS Booster or heavy ions from the Low Energy Ion Ring (LEIR) [Magistris 2008]. A section of the PS can be seen in Fig. 1.4[4].

Figure 1.4: A section of the PS during maintenance.

After the extraction from the PS, protons are sent to the Super Proton Synchrotron

[3]©2013-2015 CERN.
`http://home.web.cern.ch/about/accelerators/proton-synchrotron-booster`, 7 October 2015.
[4]©2013-2015 CERN.
`http://home.web.cern.ch/about/accelerators/proton-synchrotron`, 7 October 2015.

(SPS) where they are accelerated to 450 GeV. The SPS is the second-largest machine in CERN's accelerator complex (see Fig. 1.5[5]).

Figure 1.5: View of the SPS tunnel.

Research using SPS beams has probed the inner structure of protons, investigated nature's preference for matter over antimatter, looked for matter as it might have been in the first instants of the universe and searched for exotic forms of matter. A major highlight came in 1983 with the Nobel-prize-winning discovery of W and Z particles, with the SPS running as a proton-antiproton collider [CERN 2015].

As for the PS, since his construction the accelerator has handled many different kinds of particles: sulphur and oxygen nuclei, electrons, positrons, protons and antiprotons.

Finally the protons are transferred to the Large Hadron Collider (LHC) where they are able to reach the energy of 7 TeV. The two beams are brought into collision inside four detectors:

- A Large Ion Collider Experiment (ALICE);
- A Toroidal LHC ApparatuS (ATLAS);
- Compact Muon Solenoid (CMS);

[5]©2013-2015 CERN.
`http://home.web.cern.ch/about/accelerators/super-proton-synchrotron`, 7 October 2015.

- Large Hadron Collider beaury experiment (LHCb).

The total energy at the collision point can reach 14 TeV. After the first long shut-down (from February 2013 to March 2015), the operational beam energy at the collision point was increased from 7 TeV up to 13 TeV.

The LHC is the world's largest and most powerful particle accelerator [CERN 2015]. It consists of 27 km of superconducting magnets and accelerating structures which increase the energy of particles. The accelerated particles move, in opposite direction, inside tubes kept at a ultra-high vacuum which is of the order of 10^{-10}-10^{-11} mbar (Fig. 1.6[6]).

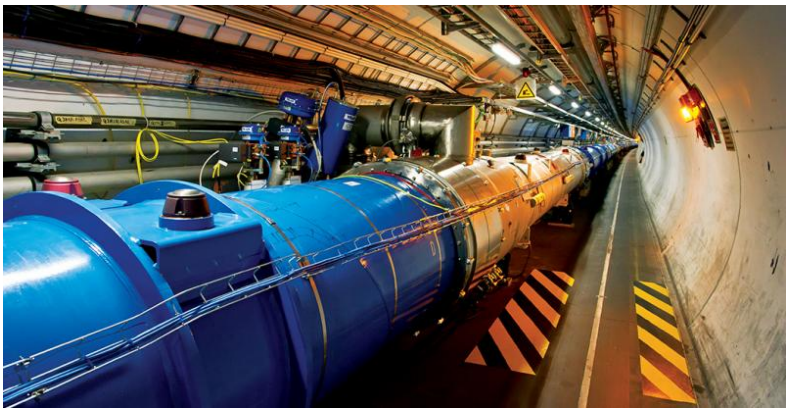


Figure 1.6: A section of the LHC, the world's largest particle accelerator.

Tab. 1.1 [Lefevre 2009] lists the most important parameters for the LHC as the luminosity[7], the nominal energy and the number of bunches (which are groups of particles in the accelerator space).

Among other facilities, the accelerator complex includes the Antiproton Decelerator (AD) providing low-energy antiprotons for the study of antimatter, the Linac 3, which is the starting point for the acceleration of heavy ions (Pb^{82+}) and the neutron Time-of-Flight

[6] ©2013-2015 CERN.
`http://home.web.cern.ch/topics/large-hadron-collider`, 7 October 2015.

[7] The luminosity $\mathcal{L}$ is the factor of proportionality between the event rate R in a collider and the interaction cross section σ_{int}: $R = \mathcal{L}\sigma_{int}$. If two bunches containing n_1 and n_2 particles collide with frequency f, the luminosity is given by:

$$\mathcal{L} = f\frac{n_1 n_2}{4\pi\sigma_x\sigma_y}$$

where σ_x and σ_y characterize the Gaussian transverse beam profiles [Chao et al. 2013].

Table 1.1: Main parameters of the LHC.

Quantity	**Number**
Circumference	26 659 m
Dipole operating temperature	-271.3 °C
Number of magnets	9593
Number of main dipoles	1232
Number of main quadrupoles	392
Number of RF cavities	8 per beam
Nominal energy, protons	7 TeV
Nominal energy, ions	2.76 TeV/u
Peak magnetic dipole field	8.33 T
Min. distance between bunches	7 m
Design luminosity	1034 $cm^{-2}s^{-1}$
No. of bunches per proton beam	2808
No. of protons per bunch (at start)	$\sim 1.15 \cdot 10^{11}$
Number of turns per second	11245
Number of collisions per second	600 million

facility (nTOF), built to study neutron-nucleus interactions.

Other major project and experiments, presently undergoing at CERN, are NA61/SHINE, studying the properties of the production of hadrons in collisions of beam particles and nuclei (such as pions[8], protons, beryllium, argon and xenon) with different targets, NA62 for the study of rare kaon[9] decays and the COmmon Muon and Proton Apparatus for Structure and Spectroscopy (COMPASS) for studying the behaviour of the interactions quarks[10]-gluons[11].

Finally, a short list of decommissioned facilities which have generated radioactive waste that is currently stored at CERN includes:

- the CERN neutrinos to Grand Sasso (CNGS), used for the study of the characteristics of neutrinos (operations ceased in 2012);

[8]The exchanged particles that carry nuclear force are called *mesons*. The lightest of the mesons, the π-meson or *pion*, is responsible for the major portion of the longer range part of the nucleon-nucleon potential. There are three types of pions with electric charges +1, 0 and -1. Their spin is 0 and their rest energies are 139.6 MeV (for $\pi^{\pm}$) and 135.0 MeV (for π^0) [Krane 1987].

[9]*Kaons* or *K mesons* are a type of meson having the property of strangeness, which is a decay caracteristic [Krane 1987].

[10]A hadron is a composite particle, made of elementary particles called *quark*, which are held together by the strong force. There exist three families of hadrons: baryons (qqq), antibaryons ($\overline{qqq}$) and mesons ($q\overline{q}$) [Hecht 2007][Mazzoldi et al. 1998]. Example of hadrons are the proton and the neutron.

[11]The force between quarks can be modeled as an exchange force, mediated by the exchange of massless particles calles *gluons* [Krane 1987].

- the Linac 1 (operated for 33 years till 1992);
- the Intersecting Storage Ring (ISR), the world's first hadron collider (operated from 1971 until 1984);
- the Large Electron-Positron collider (LEP), the largest electron-positron accelerator ever built, dismantled in 2001;
- the Low Energy Antiproton Ring (LEAR) decelerated and stored antimatter before being replaced by LEIR in 1996;
- the Synchro-Cyclotron (SC), the first accelerator built at CERN in 1957, accelerating particles at 600 MeV until 1990.

1.2 Beam losses, particle spectra and induced radioactivity

1.2.1 Beam dynamics

Beam losses are the principal cause for the activation of materials at CERN. The extent of losses is strictly monitored because even one single uncontrolled loss of a sufficiently high fraction of the beam could have major effects, such as causing a superconducting magnet to undergo a transition into a normal-conducting state or producing material damage. Nevertheless, small local and global particle losses are experienced during the normal operating conditions of accelerators. Depending on the beam intensity[12], the energy of the primary particles in a given machine and the location within the machine, total particle losses around the accelerators can span over 7 orders of magnitude [Bruce et al. 2014].

Apart from experiments where protons are voluntarily sent against fixed or mobile targets which produce a shower of secondary particles, particle loss is typically experienced by scattering with residual gaz molecules in the beam pipe and diffraction outside the acceptance region. The discussion on beam loss that follows is based on references [Magistris 2008][Wilson 2001][Wiedemann 2007][Kuhn 2013].

[12]The beam intensity is the number of particles of a kind per bunch or per year. The LHC beam intensity is $\sim 1.15 \cdot 10^{11}$ per bunch.

The design orbit of a syncrotron is a curve on which, ideally, all particles should move. In real space, most particles of the beam will deviate slightly from the design orbit. In order to limit these oscillations, focusing forces are required. Focusing and bending forces form the basis of particle motion in accelerators. The function $x(s)$ is defined to describe a transverse motion around the design orbit, commonly called "betatron oscillation". In particular, the parameter s indicates the reference trajectory of a particle and $x(s)$ the displacement function of the particle with respect to s.

Physically, the elements most frequently used in accelerators for focusing beams are quadrupole magnets. Their strength $k(s)$ (in m^{-2}) is characterized by the gradient dB_y/dx (in T/m), normalized with respect to magnetic rigidity $B(s)\rho(s)$ (in Tm):

$$k(s) = \frac{1}{B(s)\rho(s)} \frac{dB_y}{dx}. \tag{1.1}$$

The trajectory of particles in the horizontal and vertical planes is described by the solution of the homogeneous differential "Hill's equation". In the horizontal plane the equation is:

$$\ddot{x}(s) - K(s)x(s) = 0 \tag{1.2}$$

where $K(s) = [1/\rho(s)^2 - k(s)]$. The solution of Eq. 1.2 is:

$$\begin{aligned} x &= \sqrt{\beta(s)\epsilon}\ \cos[\phi(s) + \phi_0], \\ \dot{x} &= -\sqrt{\frac{\epsilon}{\beta(s)}}\ \sin[\phi(s) + \phi_0] + \left[\frac{\dot{\beta}(s)}{2}\right] \sqrt{\frac{\epsilon}{\beta(s)}}\ \cos[\phi(s) + \phi_0] \end{aligned} \tag{1.3}$$

where $\beta(s)$ is a design parameter of the accelerator, ϵ describes the area of the ellipse occupied by moving particles on the phase space and ϕ represents the phase of the quasi-harmonic motion of particles. The term $\dot{\beta}(s)/2$ is often written as $\alpha(s)$.

A typical curve describing the motion of particles in the space is the ellipse. The ideal trajectory is the centre of the ellipse and its area is a measure of how much the particles

depart from the ideal trajectory. In accelerator physics the area of the ellipse is given by:

$$Area = \pi\epsilon \tag{1.4}$$

where ϵ, called "emittance", is the product of the semi-axes of the ellipse. The emittance is usually in units of π mm mrad. The major and minor axis of the ellipse are $\sqrt{\epsilon\beta}$ and $\sqrt{\epsilon/\beta}$.

When considering beam distribution in real space, the trajectories followed by the particles have a gaussian distribution when projected on a vertical or horizontal plane. In proton machines the emittance boundary is conventionally chosen to include 87% of a gaussian beam at 2σ, where σ is the parameter describing the distribution. When accelerating protons in real space, the emittance is given by:

$$\epsilon_{protons} = \frac{(2\sigma)^2}{\beta}. \tag{1.5}$$

Finally, the "acceptance" is defined as the maximum area of the ellipse which the emittance can attain without losing particles. Machine components where the emittance is close to (e.g., septum magnets) or larger than (e.g., collimators) the acceptance, experience significant beam losses.

If $\dot{x}$ is plotted versus x, as ϕ goes from 0 to 2π, it is possible to plot an ellipse which is called "phase space ellipse", as shown in Fig. 1.7.

1.2.2 Hypothesis on beam losses

It is of common usage to express beam loss rates in terms of lost beam power. If we consider a proton accelerated at the energy E_p, we can calculate the total energy T stored in the accelerator as follows:

$$T = E_p \cdot I \cdot N_b \tag{1.6}$$

where I is the beam intensity in particles per bunch and N_b is the number of bunches. Considering for example protons moving at the maximal design energy of the LHC, the

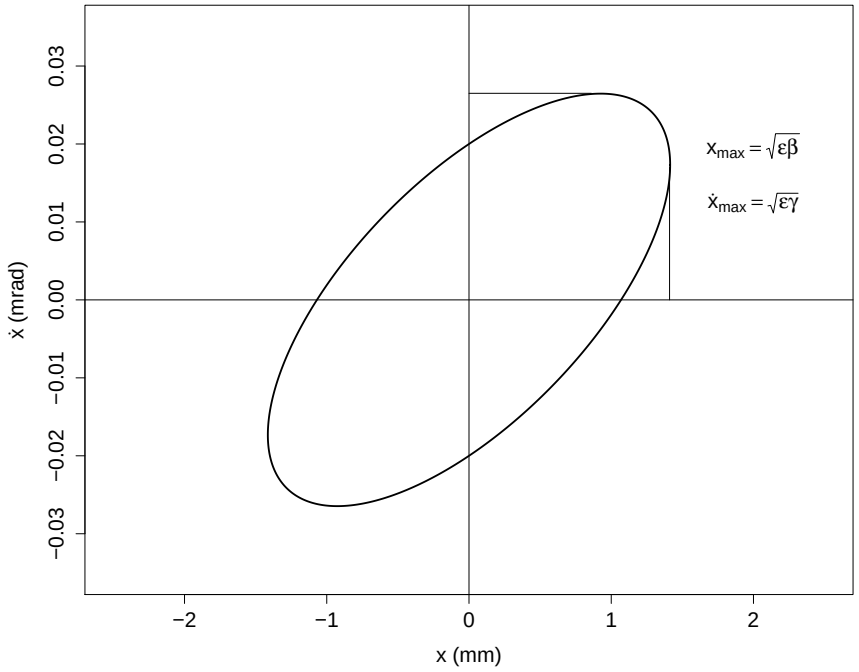


Figure 1.7: Example of horizontal phase space for a generic section of the LHC. Here $\gamma(s) = [1 + \alpha^2(s)]/\beta(s)$. Adapted from [Bruce et al. 2014] and [Wilson 2001]).

total stored energy is ∼362 MJ. To calculate the beam power we can divide the total amount of stored energy per accelerator length and per second. For the LHC, the design power is ∼13.6 kW/m.

As stated in [Plum 2012] and [Ostroumov et al.], for high-intensity linacs, losses are typically required to stay below 1 W/m. Using the design power of the LHC previously calculated, the percentage of power loss corresponding to the design limit of 1 W/m is ∼0.7%. The value found is very close to the estimated losses located at the injection of the beam from SPS to the LHC, which spans from 0.6% up to 2.47%, depending on the probability distribution of the beam [Kain].

For the activity estimation on accelerator components, the number of particle lost is needed. Using the design power limit and the energy of the principal CERN accelerators, the values of Tab. 1.2 were found. In the following chapters, the particle losses given in this table are used as guidelines to assess the relative importance of activation scenarios.

Table 1.2: Estimated particle losses per machine for the design power loss of 1 W/m.

Accelerator	**Energy/Momentum**	**Particle losses (s^{-1})**
Linac 4	160 MeV	$\sim 3.9 \cdot 10^{10}$
PS Booster	1.4 GeV	$\sim 4.5 \cdot 10^{9}$
PS	14 GeV\c	$\sim 4.5 \cdot 10^{8}$
SPS	450 GeV\c	$\sim 1.6 \cdot 10^{7}$
LHC	7 TeV	$\sim 8.9 \cdot 10^{5}$

1.2.3 Particle spectra

Due to the heterogeneity of CERN's accelerators, different particle fields exist depending on the specific machine considered but also on the location inside the tunnels. Monte Carlo codes, such as Fluka [Battistoni et al. 2006][Ferrari et al. 2005], can be used to calculate particle spectra for any specific area and its associated beam losses. However, the simulation of every possible spectrum would be impractical. Moreover such calculations, which would be very time consuming, would not necessarily lead to an improvement on predicting induced radioactivity [Magistris 2008].

To simplify tasks related to the calculation of induced radioactivity and operational radiation protection, a code called Actiwiz was developed at CERN [Theis and Vincke 2012]. Actiwiz defines a limited number of simplified scenarios which are common at CERN. Using these simplified scenarios, which rely on a set of activation parameters like energy, position, irradiation time and decay time, the user can calculate complete radionuclide inventories, evaluate the contribution of each radionuclide to predefined hazard factors and identify chemical elements generating each produced isotope. A detailed description of Actiwiz will be given in Ch. 2.

In order to obtain activity estimations and radionuclide inventories, Actiwiz relies on an extensive number of spectra calculated using the Fluka code [Battistoni et al. 2006] [Ferrari et al. 2005]. For induced radioactivity studies in hadron accelerators, the particles that are of major interest are neutrons, protons and pions (positive and negative). Depending on the location inside the tunnel, activation occurs mainly by interaction with low energy neutrons (such is the case for the locations close to the tunnel walls or behind concrete shielding) and highly energetic hadrons (location close to the beam area). In Figs. 1.8 and

1.9 we show two examples of spectra corresponding respectively to the location close to the wall shielding and beam impact for the LHC.

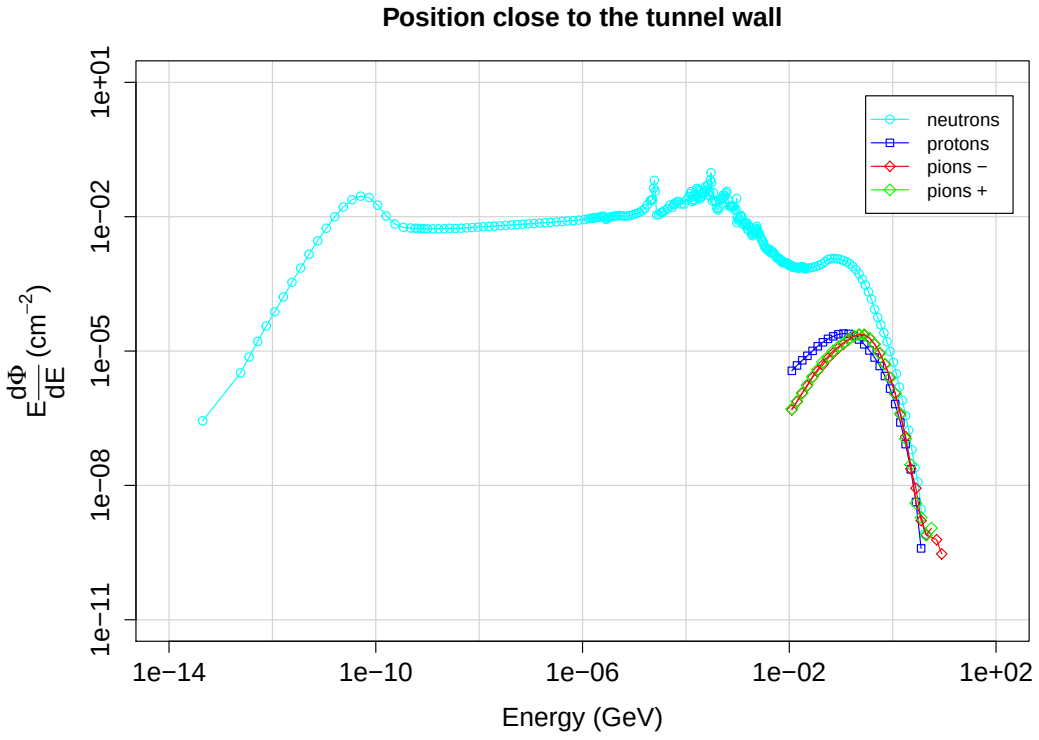


Figure 1.8: Particle fluence generated by a current of one proton per second inside the LHC, when the position is close to the concrete tunnel wall (the data used to build the spectrum is a courtesy of H. Vincke and C. Theis, CERN).

In particular, in Fig. 1.8 we can see the fluence spectrum of secondary particles generated when a current of one proton per second is lost in the LHC and the position is close to the concrete tunnel shielding. It is easy to observe that the fluence spectrum is dominated by the neutron component (colored in blue).

In Fig. 1.9 we can observe the fluence of secondary particles generated by 7 TeV protons at the beam impact area. In this case we can see a higher contribution of highly energetic hadrons to the spectrum.

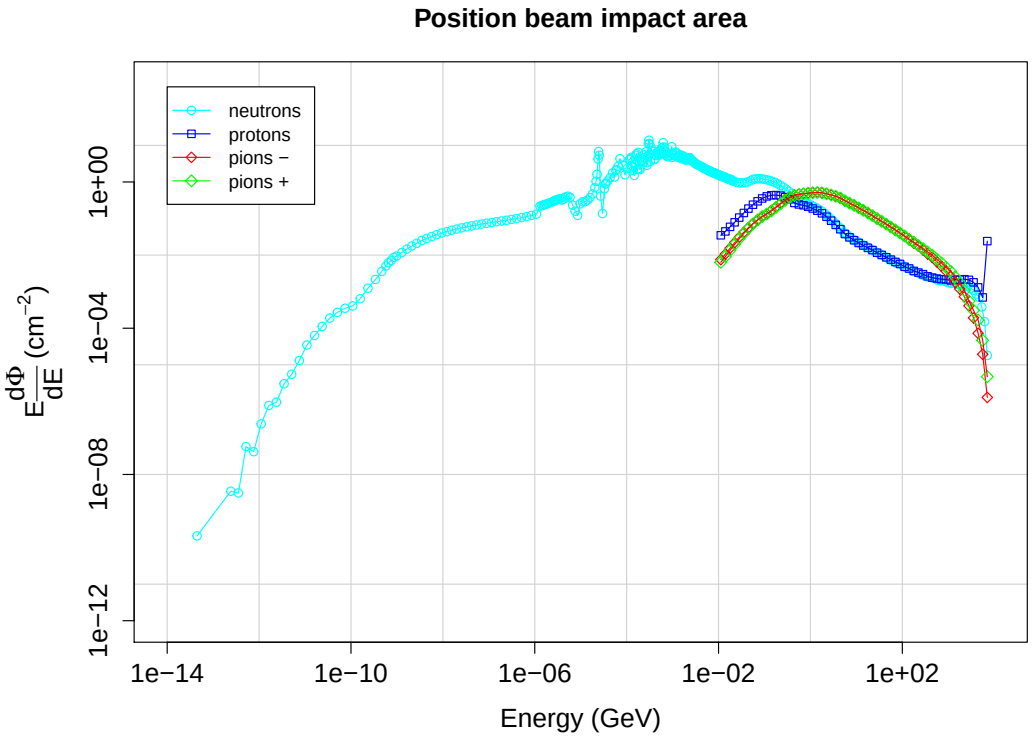


Figure 1.9: Particle fluence generated by a current of one proton per second inside the LHC for the beam impact area (the data used to build the spectrum is a courtesy of H. Vincke and C. Theis, CERN).

1.2.4 Induced radioactivity

The activation process can be described as the sum of nuclear reactions which made a material radioactive. The beam particles that are lost can interact with the structures surrounding the accelerators. Depending on the interaction conditions, materials can be activated. The activation formula presented in this section is based on references [Barbier 1969] and [Froeschl et al. 2014].

The production rate of isotope r from element e for a loss of one primary particle per second is given by:

$$P_{re} = \frac{\mathcal{N}_A}{\mathcal{M}_e} \sum_{i=p,n,\gamma,\pi^+,\pi^-} \int \Phi_i(E)\sigma_{i,e,r}(E)dE \tag{1.7}$$

where $\mathcal{N}_A$ is the Avogadro's number, $\mathcal{M}_e$ is the atomic weight of the element e, the sum is

extended over protons (p), neutrons (n), photons (γ) and pions (π^+, π^-), $\Phi(E)$ denotes the radiation fluence for the various secondary particles generated by one primary particle per second and $\sigma_{i,e,r}(E)$ is an abudance weighted average of the cross sections of each isotope of element e.

The time evolution T_{br} of the specific activity of isotope b, for an irradiation time t_i followed by a decay time t_c is given by:

$$T_{br}(t_i, t_c) = \sum_{c, r \to b} \sum_{m=1}^{j_c} \frac{c_m^c}{\lambda_m^c} e^{-\lambda_m^c t_c} \left(1 - e^{-\lambda_m^c t_i}\right) \tag{1.8}$$

where c runs over all decay chains starting from isotope r leading to isotope b and j_c is the number of isotopes in a given decay chain c. λ_m^c denotes the total decay rate of the mth isotope in the decay chain c. $\tilde{\lambda}_m^c$ denotes the partial decay rate of the mth isotope in the given decay chain c and the coefficient c_m^c is the Bateman coefficient of the mth isotope in decay chain c given by:

$$c_m^c = \frac{\prod_{i=1}^{j_c} \tilde{\lambda}_i^c}{\prod_{\substack{i=1 \\ i \neq m}}^{j_c} (\lambda_i^c - \lambda_m^c)}. \tag{1.9}$$

The specific activity of isotope b induced by a loss rate of one primary particle per second is then given by:

$$a_b = \sum_r \sum_e T_{br} P_{re} m_e \tag{1.10}$$

where m_e is the weight fraction for the element e, r represents all the isotopes directly produced and e the generic element of the material.

In Actiwiz [Theis and Vincke 2012], Eq. 1.7 is evaluated by Fluka [Battistoni et al. 2006][Ferrari et al. 2005] and Eq. 1.8 is solved analytically.

1.3 Radioactive waste and disposal pathways

Radioactive waste produced at CERN is disposed of in France and Switzerland depending on the disposal pathways existing in the host states. The regulatory framework covering the elimination of radioactive waste at CERN is described by the so-called "Tripartite Agreement" [Acc 2011] and by the Swiss "Ordonnance sur la Radioprotection" [ord 1994]. The Tripartite Agreement was ratified by CERN, French government and Swiss Federal Council the 15 November 2010 to regulate the protection against radiation and the safety of facilities of the European Organization for Nuclear Research.

With respect to the disposal of radioactive waste and as required by the Tripartite Agreement, CERN writes and updates a waste study covering and inventorying the waste produced in each facility of the Organization. The waste study proposes the best disposal pathway for a waste, taking into account the available information and the need of a fair share between host states in terms of quantity, activity and toxicity.

The following sections describe the disposal pathways for radioactive waste at CERN at the time of writing. In particular, the elimination of VLLW in France[13] and the clearance practices in Switzerland will be described. We also list a short summary of radioactive waste presently stored at CERN.

1.3.1 VLLW disposal pathway

To test the acceptability of a radioactive waste package in the final repository as a VLLW, the French National Agency for Radioactive Waste Management (ANDRA)[14] has defined a factor called IRAS[15] as follows:

$$IRAS = \sum_i \frac{a_i}{AL_i} \tag{1.11}$$

where a_i is the specific activity of the radionuclide i (in Bq/g) in the mass of the given waste and AL_i is a coefficient expressing the level of radiotoxicity of the radionuclide i.

[13] In France VLLW is identified as waste "Très Faiblement Actifs" (TFA).
[14] `http://www.andra.fr/`, 13 October 2015.
[15] In French IRAS stands for "Indice Radiologique d'Acceptabilité en Stockage".

This coefficient is given by:

$$AL_i = 10^{Class_i} \tag{1.12}$$

where the $Class$ of the radionuclide i states its level of radiotoxicity. The classes vary between 0 and 3, with class 3 being the class of the radionuclides with lower toxicity. The acceptance requirements for VLLW waste and a detailed explanation of the IRAS can be found in references [ANDRA 2013a] and [ANDRA 2013b].

Table 1.3 shows a list of nuclides, with related properties, which are of interest for the radiological characterization of metallic VLLW, produced in hadron accelerators.

The Declaration Threshold (DT) given in this table is the activity level above which a radionuclide must be declared and used for the computation of IRAS.

The acceptability of a collection of packages of radioactive waste is tested introducing a weighted IRAS as follows:

$$\langle IRAS \rangle = \frac{\sum\limits_k M_k \cdot IRAS_k}{\sum\limits_k M_k} \tag{1.13}$$

where M_k is the weight of the k-package of the batch and $IRAS_k$ is the IRAS of the k-package calculated by the Eq. 1.11.

Necessary conditions for a batch of waste to meet the acceptance criteria are $\langle IRAS \rangle < 1$ and IRAS< 10 for each waste package of the batch.

1.3.2 Clearance

Clearance from regulatory control of very-low-level radioactive materials and waste produced at CERN can be performed in Switzerland if the radiological characteristics of the objects considered satisfy the criteria outlined in the reference documents [ord 1994][ifs 2009]. The Swiss Federal Office of Public Health (FOPH) verifies the measurements performed by CERN at this aim. In this subsection we shortly describe technical information related to solid waste elimination for clearance purposes.

To understand the technical requirements for clearance of materials, some preliminary

Table 1.3: Half-life ($T_{1/2}$), VLLW class and Declaration Threshold (DT) of common radionuclides of activated metals in hadron accelerators [ANDRA 2013a]. Columns 4 and 5 give reference values for clearance in Switzerland [ord 1994]. LE is the so-called Exemption Limit and CS is the Surface Contamination.

Nuclide	**$T_{1/2}$ (y)**	**Class**	**DT (Bq/g)**	**LE (Bq/kg)**	**CS (Bq/cm^2)**
H-3	12.312	3	1	$2 \cdot 10^5$	1000
C-14	5730	3	0.1	$2 \cdot 10^4$	30
Na-22	2.603	1	0.1	3000	3
Al-26	$7.16 \cdot 10^5$	1	0.1	3000	3
Cl-36	$3.01 \cdot 10^5$	3	0.01	10^4	3
Ar-39	269	3	10	-	-
Ca-41	$1.03 \cdot 10^5$	3	0.01	$3 \cdot 10^4$	300
Ti-44	58.9	1	0.1	2000	30
V-49	0.903	3	10	$6 \cdot 10^5$	100
Mn-54	0.855	1	0.1	10^4	100
Fe-55	2.73	3	10	$3 \cdot 10^4$	300
Co-57	0.744	2	1	$5 \cdot 10^4$	100
Co-60	5.2711	1	0.1	1000	3
Ni-63	100	3	10	$7 \cdot 10^4$	1000
Zn-65	0.669	1	0.1	3000	30
Ge-68	0.741	1	0.1	8000	3
Mo-93	4000	3	0.01	4000	300
Nb-93m	16.12	3	10	$8 \cdot 10^4$	1000
Nb-94	$2 \cdot 10^4$	1	0.1	6000	3
Tc-99	$2.1 \cdot 10^5$	3	0.01	1000	3
Rh-101	3.3	2	0.1	$2 \cdot 10^4$	10
Rh-102	2.9	1	0.1	4000	30
Ag-108m	437.7	1	0.00025	4000	30
Ag-110m	0.684	1	0.1	4000	10
Sn-119m	0.802	3	10	$3 \cdot 10^6$	300
Sn-121m	50	3	10	$3 \cdot 10^4$	30
Sb-125	2.7586	1	0.1	9000	10
Gd-148	74.6	1	0.1	200	1
Hg-194	520	1	0.1	200	3
Bi-207	31.55	1	0.1	8000	30

concepts must be introduced. In particular, the following terms will be described: Surface Contamination (CS), dose rate ($\dot{D}$) and Exemption Limit (LE). Numerical values for the limits to use in clearance of metallic waste activated in hadron accelerators can be found in Tab. 1.3.

The surface contamination is given by the sum of the non fixed activity (which could be removed from a surface by wiping or washing) and the fixed activity on a surface which could be removed during future use. This quantity is often expressed in Bq/cm^2.

The equivalent dose rate is an operational quantity to estimate the exposure of a worker

in a given environment. This quantity is measurable (for example using a portable dose-rate meter) and is used to extrapolate potential exposure values, like for example the total dose integrated when performing a task. The unit for equivalent dose rate are Sv/h; common subunits are mSv/h and μSv/h.

The LE is an activity threshold (in Bq/g or Bq for LE absolute which is the maximum activity allowed for a specified item and a given radionuclide) below which the rules given in reference [ord 1994] are not applicable. More precisely, the ingestion of 1 kg of a substance having a specific activity of 1 LE, leads to a maximum Committed Effective Dose $E(\tau)$[16] of 10 μSv.

Solid radioactive waste is considered as inert and can be treated as conventional waste (after release from regulatory control) if the three following requirements are fulfilled simultaneously:

- the surface contamination is below or equal to the corresponding CS value given in reference [ord 1994]; examples of CS are given in Tab. 1.3;
- the equivalent dose rate at 10 cm from the surface is below or equal to 0.1 μSv/h;
- the specific activity is below or equal to 1 LE and the total activity of waste eliminated in a month is below 100 LE.

1.3.3 Radioactive waste from CERN accelerators

The production of radioactive material in an accelerator environment is a continuous process. Normally items of small dimensions are replaced during short-time periods of maintenance while bigger items (like magnets) are only delivered to the radioactive waste storage during longer shut-down periods.

From the point of view of material activation, it is during the machine operation that the accelerator components become radioactive. However, it is only at the time of dismantling

[16]Following an intake by ingestion or inhalation, some radionuclides persist in the body and irradiate the various tissues for many years. The total radiation dose in such cases depends on the half-life of the radionuclide, its distribution in the body, and the rate at which it is expelled from the body. The resulting total effective dose delivered over a lifetime (the commitment period is taken to be 50 years for adults and to age 70 years for children) is called the Committed Effective Dose.

that an item is regarded to as radioactive waste [Magistris 2008].

A constant amount of material (of the order of few hundreds cubic meters per year) is sent to the treatment and storage centre during operation of the accelerators but the largest fraction of radioactive waste arises from special decommissioning and dismantling projects.

The major projects that led to the present amount of waste stored at CERN can be summarized as follows:

- the SC improvement program, with important modifications of the SC vacuum chamber, the radio-frequency system and the ISOLDE facility connected to it in 1973;
- the Antiproton Collector project which required the complete dismantling and modification of the Antiproton Accumulator target area in 1986;
- the cleaning operation of the neutrino cavern of the SPS in 1992;
- dismantling of the magnets of the ISR in 1993;
- the dismantling of LEP in 2001;
- the dismantling of West Area Neutrino Facility (WANF) in 2009;
- the dismantling of the SC complex and the ISOLDE 1 experience in 2010;
- the partial dismantling of CNGS in 2013;
- the Long Shutdown 1 in 2013-2014 to increase the LHC beam energy.

A significant part of the radioactive waste collected from dismantling facilities are metals. Waste metals consist primarily of steel, copper and aluminium. For this reason the work performed in the present thesis focuses on the radiological characterization of this family of materials. Nevertheless, the method that we describe in the following chapters is presently tested on other materials such as plastic. We describe the general outline of the radiological characterization philosophy in the last section of the present chapter.

1.4 Radiological characterization

To characterize a radioactive waste means to accurately assess the physical, chemical and radiological characteristics of the waste, with a view to verify if the waste meets the acceptance criteria established by the disposal facility. Reference [IAEA 2007] gives a detailed description of the strategies and methodologies that can be used to characterize a radioactive waste.

The waste characterization process depends on many factors such as:

- the type of waste (for instance, high or low level waste, liquid or solid);
- the regulatory regime (clearance or radioactive waste disposal);
- the available knowledge on the activation process;
- the parameters to be measured.

Because of the factors cited, it is not possible to identify a single characterization procedure which is applicable to each waste type. It is then necessary to fully understand the waste acceptance criteria and the safety assessment of the disposal concept in order to elaborate a strategy for the characterization actions. Such general strategy can then be applied to the establishment of specific characterization procedures for a given type of waste.

A list of the key parameters to assess during the characterization process includes:

- waste volume and quantity;
- nuclide inventory;
- activity;
- chemical properties;
- biological properties;
- production process.

The present thesis focuses on the radiological aspects of the waste characterization process. It then discusses a structured methodology to evaluate nuclide inventories and specific activities for VLLW produced at CERN.

We propose a method tailored for VLLW metals which represent the vaste majority of radioactive waste currently stored at CERN.

Even if the method is presently used for the radiological characterization of legacy waste, there are no theoretical limitations to apply it to the characterization of new waste. This section, after a short description of "new" and "legacy" waste, ends with the flow diagram of the new radiological characterization process.

1.4.1 New and legacy waste

For the purpose of waste characterization it is useful to distinguish "new" from "legacy" or "historical" waste. Waste is considered as "new" if generated when a traceable characterization programme is in place. For this category of waste the history is known, the characterization is complete and the process knowledge is appropriately stored and managed.

Legacy waste is generated before a complete traceable characterization program was in place. The present thesis proposes specific procedures for the characterization of legacy, VLLW metallic waste. For its characteristics, legacy waste often belongs to the category of "complex and variable waste" as defined in [IAEA 2007]. This means that the identification of the nuclide inventory is often difficult, unstable and that the process for the waste characterization is expensive. For instance, mixed non traced waste is particularly difficult to sample and characterize.

We implemented statistical methods and coupled them with dedicated measurements to create a robust characterization program for this family of waste. Moreover, the radiological characterization of waste produced at particle accelerators is not standardized, such as for the case of waste produced in nuclear power plants. We therefore adapted the known standards and techniques used in the nuclear industry to the needs of research laboratories and accelerating facilities.

1.4.2 The radiological characterization workflow

The radiological characterization workflow is presented in Fig. 1.10. The main steps are identified with a letter (from A to H) and will be discussed in detail in the following chapters.

The radiological characterization of a waste starts with the definition of the perimeter of the waste population (step A). For instance, a waste population can be made of all the legacy metallic waste that can be compacted or all cables activated at CERN. For practical reasons each population (which can add up to thousands of tons) is split into sub-populations which are commonly called "batches". When the perimeter of a population is identified and the boundary of each sub-population is clear, we start collecting information that is important for the successive steps. We are primarily interested in knowing the nature of the activated materials (for instance copper, concrete or steel), their volume, the origin of the waste, the accelerator in which the material was installed and the radiological history. The radiological history includes the irradiation time (or time during which the waste was irradiated) and the decay time (time elapsed since the waste was last irradiated).

For legacy waste it is often very difficult to obtain such information and working hypotheses are necessary to evaluate the preliminary radionuclide inventory (step B). For instance, if the decay time is not known, it is necessary to consider a set of potential decay times when we simulate the radionuclide inventory. This is commonly done making the decay time vary from 1 up to 30 or 40 years in the simulation framework.

After the necessary data is collected we inject them within analytical or Monte Carlo codes that are used to make predictions about the radionuclides produced when a material is exposed at CERN's irradiation scenarios. A scenario is given by the unique combination of a chemical composition, the energy of an accelerator, the position of the material within the accelerator, an irradiation time and a decay time. The complete list of produced radionuclides is called "radionuclide inventory".

Each radionuclide can be classified as either ETM, DTM or ITM, as previously defined. Easily measurable radionuclides (ETMs) can be quantified via non-destructive techniques from outside a waste package, such as γ-ray spectrometry. DTM radionuclides are difficult to

measure because the quantification of their activity requires complex destructive techniques. Finally, ITM radionuclides are similar to DTMs but their activity is generally so low that even non-destructive methods are not able to quantify their concentration. Their activity is instead evaluated using analytical or Monte Carlo codes. ETM radionuclides whose activity is correlated to the activity of DTMs or ITMs are called key nuclides (KN).

When the radionuclide inventory is established, we can apply various methods to estimate the activity of each radionuclide in the batch of waste under consideration. The process developed at CERN requires a mix of direct measurements (step C) and calculations (steps D and E) to evaluate the activity. Direct measurements consists of systematic γ-ray spectrometries performed on each final waste package. The calculations used to evaluate the activity of DTMs and ITMs radionuclides include the so-called "scaling factor" method (SF), the "mean activity method" (MA) and the "correlation factor" method (CF).

The SF method lies upon an experimental correlation between the activity of a KN with a DTM. If this experimental correlation exists and it's proved over an adequate sample, the activity of DTM radionuclides can be obtained by multiplying the activity of the KN by the scaling factor.

Similarly, if a theoretical correlation between a KN and an ITM can be established via analytical calculations or Monte Carlo simulations, the activity of the ITM can be obtained via the multiplication of the activity of the KN and the established CF.

The activity obtained in the previous steps is finally used to evaluate the IRAS of each waste package and the IRAS of the batch (step H). The uncertainty of the activities is quantified (step G) and the IRAS is compared to the limits to evaluate the waste acceptability at the disposal facility.

Chapter 2 describes in detail step B of the characterization workflow. In Ch. 3 we present the methodologies that we use to quantify the activity of ETM, DTM and ITM radionuclides (steps C to F). The uncertainty quantification is outlined in Ch. 4 and examples of application of the methods introduced are finally discussed in Ch. 5.

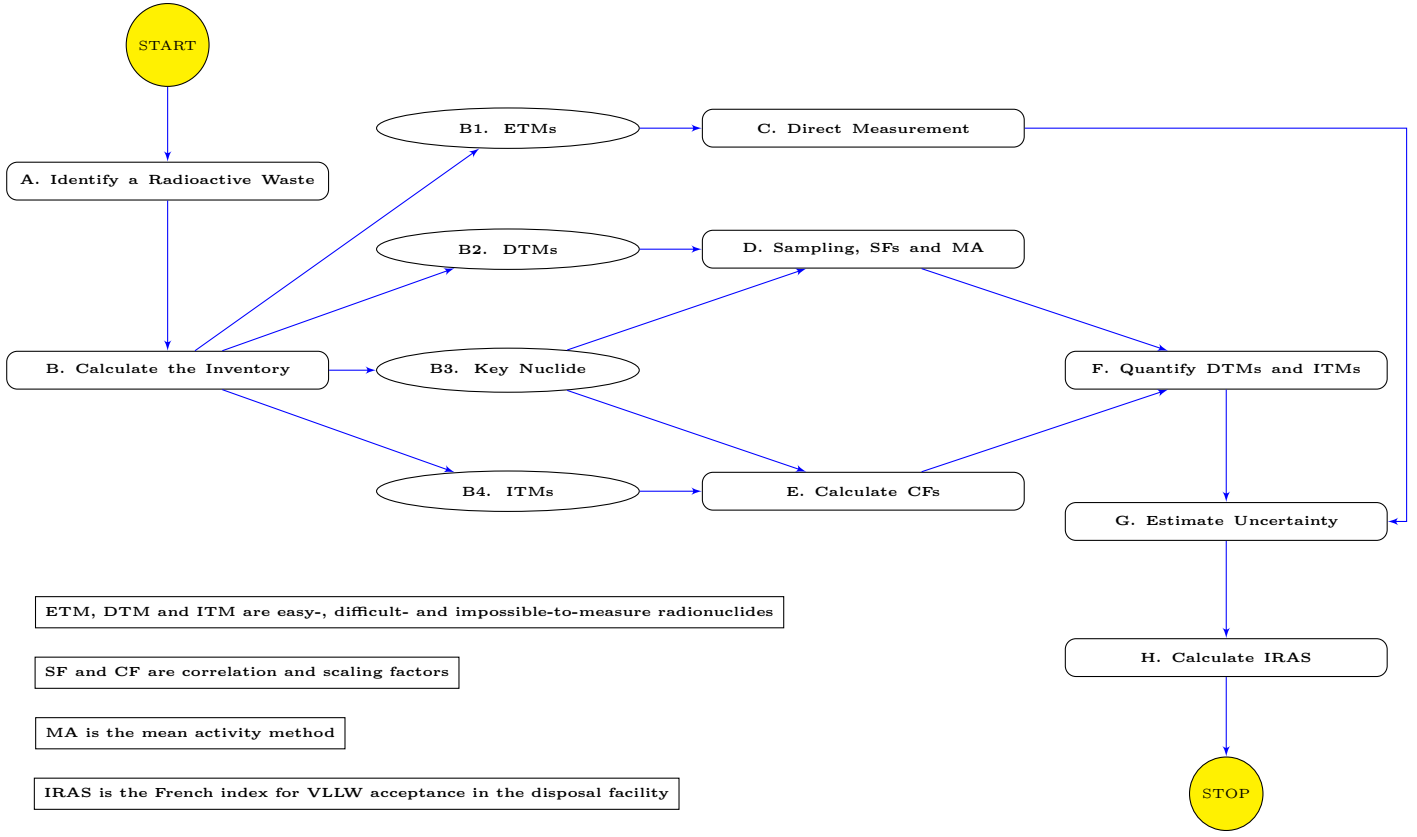


Figure 1.10: Radiological characterization process developed at CERN for VLLW metallic waste.

Summary

In the present chapter we introduced the complex of CERN's accelerators and discussed the main reasons leading to material activation and radioactive waste production. We discussed the disposal pathways presently available at CERN and introduced the radiological quantity IRAS, which must be calculated to evaluate the acceptability of a waste into disposal facilities.

The characterization method that we developed targeted legacy metallic waste because they represent the majority of the waste currently stored at CERN.

The characterization method consists of the following sequence of steps according to Fig. 1.10:

A. Identification of a radioactive waste population

B. Evaluation of the radionuclide inventory through analytical calculations or Monte Carlo simulations

C. Quantification of the activity of ETMs via direct measurement

D. Applicability tests for scaling factors and mean activity method

E. Quantification of correlation factors

F. Quantification of DTMs and ITMs activity

G. Estimation of uncertainty

H. Evaluation of the IRAS for the acceptability of the waste in the disposal facility.

Chapter 2

Evaluation of the radionuclide inventory

Remember that all models are wrong; the practical question is how wrong do they have to be to not be useful.

George Box
Empirical Model-Building and Response Surfaces.

Introduction

The radionuclide inventory is the list of radionuclides in a waste, produced with an activity above the Declaration Threshold. For VLLW the establishment of this list is of outmost importance because it allows national agencies of waste management (ANDRA in France and FOPH in Switzerland) to evaluate the compliance with the safety requirements of the disposal site. This evaluation includes the impact on the population and the environment of potential accidents in which such radionuclides are released.

The steps of the radiological characterization process that we developed and that are related to the establishment of the radionuclide inventory in VLLW produced at CERN are showed in Fig. 2.1. The complete flow diagram can be seen in Ch. 1, Fig. 1.10. We recall here that ETM, DTM and ITM stand for easy-, difficult- and impossible-to-measure radionuclides, according to the definitions given in Ch. 1. We would like also to point out that the majority of waste at CERN is produced by means of activation (the contamination

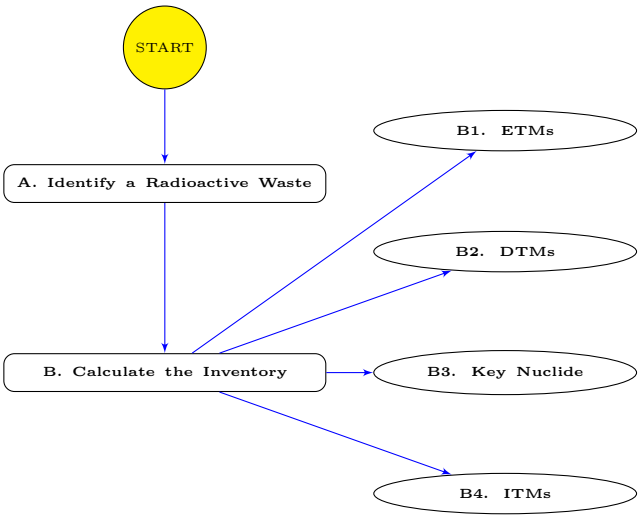


Figure 2.1: First stages of the radiological characterization process. ETMs, DTMs and ITMs stand for easy-, difficult- and impossible-to-measure radionuclides.

is limited to experimental areas such as nTOF or ISOLDE). Therefore, in the following chapters we will use the term "waste" as synonym of "activated waste" (for more details see Sec. 1.2.4).

Every disposal campaign starts with the identification of the boundaries of a waste family (step A of the process). A number of criteria can be chosen to identify those boundaries including similarity on material composition, dismantling date, origin, similarity of radiological history or treatment requirements. Figs. 2.2 and 2.3 show two examples of waste families considered for disposal at CERN in 2016.

A number of theoretical activation methods have been developed to evaluate the radioactivity of waste generated at nuclear reactors. Even if radioactive waste produced at nuclear reactors is different from the one generated at CERN (we can cite, for instance, the very limited use at CERN of nuclear materials such as uranium and the limited presence of contamination), the basic principles of the activation methods are still applicable. A comprehensive reference for this topic is the ISO standard 16966 [ISO 2013].

According to reference [ISO 2013] we can distinguish between the so-called "point method" and "range methods".

The point method is a calculation technique, generally based on the use of analytical

Figure 2.2: An example of waste family produced and stored at CERN. These metallic elements are accelerating cavities of the LEP machine.

Figure 2.3: An example of activated cables produced at CERN, presently being disposed of as VLLW.

or Monte Carlo codes, that allows us to evaluate the activation of a single item or of a point of an object with high accuracy because it lies on precise data about the material of the item, the irradiation conditions, the particle fluence rate and its position. With the point method we can estimate the specific activity of the point considered with a low uncertainty since the input conditions are well known. The range methods apply to a collection of components of similar type and irradiation history but taking into account stochastic variation of the input parameters like the chemical composition, the particle fluence rate and the radiological history. The range methods give the distribution and the average activity of a collection of activated items.

Among the range methods, we consider here the so-called "correlation method" and "distribution evaluation method" that we have adapted to characterize waste produced at particle accelerators.

The correlation method evaluates the activity of radionuclides which are produced with a specific range of elemental composition, radiological history, particle fluence rate and energy. If a correlation between a given radionuclide and the key nuclide KN (major γ-emitter) exists, the activity of the given radionuclide is obtained multiplying the correlation factor by the activity of the KN. The correlation factor is a coefficient obtained via simulations, linking the activity of ETM and DTM or ITM radionuclides. With this method a correction with time must be implemented because of the different half-life of the radionuclides within the waste.

The distribution evaluation method is similar to the correlation method but extends the variation of the input parameters to the location and, consequently, the particle spectra. The distribution evaluation method allows us to build activity distributions for the produced radionuclides. The appropriate average content estimator for the specific activity can finally be chosen according to the underlying distribution.

The process for performing activation calculations in order to estimate the concentration of radionuclides in waste involves multiple steps [ISO 2013]. For the very nature of the waste already stored at CERN we will focus on the range methods. We use the point method for the rare cases in which the input parameters are well known (for instance, the targets built for facilities like ISOLDE and nTOF).

Once the method is chosen it is crucial to select and determine the input parameters. The calculations are finally carried out and the raw data are processed to evaluate the correlation factors, the distributions or the conversion factors according to the methodology adopted.

In the following sections we describe in detail the characteristics of the input parameters for the activation calculations carried out to characterize VLLW at CERN. In particular, we will focus on the definition of a so-called scenario **S** representing the realisations of a set of random variables including the chemical composition (**CC**), the beam energy (E), the position inside the tunnel (P), the irradiation (t_i) and the decay (t_c) times. The scenario is used as input parameter for the calculations allowing us to estimate the radionuclide inventory and to identify ETM and the KN (sub-steps B.1 and B.3), DTM (sub-step B.2) and ITM (sub-step B.4) radionuclides within a waste population.

From a mathematical point of view **S** is a mixed multivariate random variable (or vector). It is mixed because the random variables can be both discrete and continuous. It is multivariate because it includes multiple variables. A realization of the object scenario is a vector of dimension n with values in $\mathbb{R}^n$. The dimension n depends on the number of elements of the chemical composition involved while the other components of **S** have each dimension 1. In practice, the chemical composition is also a multivariate random vector and its dimension depends on the number of chemical elements it is made of.

We can write the scenario **S** as follows:

$$\mathbf{S} = (\mathbf{CC}, E, P, t_i, t_c). \tag{2.1}$$

Once a realization of **S** is available it can be used as an input parameter for numerical experiments. In practice, each realization can be used as the input of either a Monte Carlo simulation or an analytical calculation. The production of a large number of realizations can be considered to be a robust and fast method to investigate the activation of materials and it is a valid substitute of long and expensive experiments, the results of which can be difficult to generalize.

After a short introduction of the simulation tools used in the present thesis, the following

sections describe in detail the components of Eq. 2.1 and the methodology adopted to generate the realizations of the mixed multivariate random vector. We will also show detailed practical examples of the methodology previously described.

2.1 Calculation and simulation codes

The present section introduces the Monte Carlo methods and the calculation tools used to obtain approximate solutions to activation of material for radiological characterization purposes. The Monte Carlo code Fluka [Battistoni et al. 2006][Ferrari et al. 2005] is introduced together with an overview of the physical models used. The calculation code Actiwiz [Theis and Vincke 2012], a tool to estimate hazard factors related to irradiated materials at CERN, is also presented with a special focus on radiological studies for waste elimination, including IRAS calculation, estimation of radiological inventories and evaluation of correlation factors.

2.1.1 Monte Carlo methods and the Fluka code

The analytical formulation of physical problems, like particles transport or radiation interaction with matter, is complicated and in the most complex cases impossible. As an alternative, Monte Carlo methods using random sampling can be used to calculate approximate solutions of their analytical equations. Heavy mathematics can be avoided entirely by mean of simulations [Dunn and Shultis 2011][La Torre 2014].

Monte Carlo techniques propose a solution of the transport equations without solving the mathematics describing the problem. A virtual experiment, similar to the real one, is solved instead. Each physical event is simulated and the particle stories are individually tracked. For an accurate reproduction of the physical phenomenon, a description of each region with its geometrical boundaries and material composition is required. This information is often provided through an input file, while an adequate description of the cross sections for each isotope is contained in the libraries used by the codes [La Torre 2014].

Once each event that may happen to a particle is defined in terms of its associated

probability, any required macroscopic physical quantity can be scored at any point of the mathematical experiment. Due to the statistical nature of the Monte Carlo method, the result of the simulation approaches the true value of the real physics quantity only if the number of events treated approaches infinity. Since a real calculation does not use an infinite number of particles, results obtained by simulations must be considered with some judgement [Kroese et al. 2011].

The significance of the results found using Monte Carlo simulations may be improved by using biasing techniques, e.g. the "Splitting method" and the "Russian roulette" [La Torre 2014][Kroese et al. 2011][Ferrari et al. 2005]. With these techniques a different weight is given to the trajectories and the regions that are supposed to give the greatest contribution to the physical quantity scored.

Among the Monte Carlo codes commercially available, the Fluka transport code [Battistoni et al. 2006][Ferrari et al. 2005] is particularly tailored for CERN needs. Fluka is a general purpose Monte Carlo code developed by the Italian Nuclear Physics Institute in collaboration with CERN and Houston University. The code allows the user to simulate particles transport and their interactions with matter. In particular, Fluka simulates the effects of the interactions of more than 60 different particles, such as photons and electrons (with energies ranging from 1 keV to thousands of TeV), neutrinos, muons and hadrons (up to 20 TeV) and heavy ions.

In the context of the present study, the interest of the code lies on the possibility to predict the induced radioactivity on materials during hadronic and electronic cascades. The code is also able to calculate the time evolution of the radiological inventory with the exact analytic implementation of Bateman's equations [Bateman 1910]. The CERF experiment [Mitaroff and Silari 2002] shows that most of the radionuclides produced inside accelerators are predicted with errors below 20% for a 95% confidence level [Brugger et al. 2007].

Several theoretical models are employed in Fluka for the transport of the different groups of particles and for different energy ranges. Isotope production by hadronic interactions (except for low-energy neutrons) is described by a Generalized Intra-Nuclear Cascade model called PEANUT (Pre-Equilibrium Approach to Nuclear Thermalization) [Ferrari

et al. 2005]. The Glauber-Gribov approach [Arsene et al. 2007] together with the Dual Parton Model [Ranft 1995] is used to describe the high-energy (above several GeV) primary interactions with target nucleons. Various complementary mechanisms are also implemented for evaporation, fragmentation, fission and de-excitation by γ-emission. A detailed bibliography on these topics can be found in [Ferrari et al. 2005].

Finally, the transport of neutrons with energies below 20 MeV is performed by a multi-group algorithm based on evaluated cross section data (ENDF/B, JEF, JENDL, etc.) binned into 260 energy groups, 31 of which are in the thermal energy region [La Torre 2014].

2.1.2 The Fluka input file

To generate a Fluka input file, the most important elements to define are:

- the source term;
- the geometry;
- the materials (chemical composition and density);
- the desired outputs.

In the context of the present study the source term is known because it is defined by the radiological environment. A radiological environment coincides here with a machine of the CERN's accelerator complex (see Ch. 1). The calculations performed for the validation of the present thesis considers mainly protons as source term. For the LEP machine, today dismantled, the appropriate source term consists of electrons.

Geometries of accelerators and tunnels are also available because of the great number of studies that have been carried out for dosimetry, shielding and activation of materials but also because of the high number of calculations performed during the accelerators design phase.

In Fluka, geometries are generated following the rules of combinatorial geometry (each region is defined by union, intersection, exclusion and complement of surfaces previously

defined). A material is then associated with each region. Particles can be followed region by region or defining an environment which is independent from regions [Magistris 2008][Zaffora 2014].

The combinatorial geometry is based on the two concepts of body and region. Bodies are defined as convex portions of space (i.e. finite portion of space completely delimited by surfaces of first or second degree, infinite cylinders and planes). Complex objects, called regions, are defined combining bodies by using Boolean operators. Each region is not necessarily simply connected, since it can be made of two or more non contiguous parts, but must be of homogeneous material composition. All the regions are surrounded by an infinitely absorbing material (called "blackhole"), which absorbs all the escaping particles.

A repetition capability (called "lattice") is also available to avoid the multiple description of repetitive structures. Various visualization tools, such as Flair[1], have also been integrated into Fluka in order to simplify the geometry building process and the debugging phase [La Torre 2014].

When simulating induced radioactivity using Fluka, at least two possibilities are available. A first possibility consists in evaluating the production rate per primary particle followed by off-line calculation during the post-processing phase. A second method consists of using Fluka to calculate the specific activity accumulated after a given irradiation cycle [Magistris 2008].

Fluka is a powerful tool since it can simulate very complex systems and is nowadays essential for accelerator design, activation studies and dosimetry or space applications. For the characterization of legacy waste we are however confronted with large uncertainty associated with the input parameters, including the radiological history of a waste item or even its elemental composition. In such cases, a Monte Carlo code like Fluka can be cumbersome to use because we should recreate a new simulation for each variation in the input parameter. This often implies heavy modification of the input file and a long running time to obtain the desired precision.

As it will be discussed in the following section, a tool called Actiwiz [Theis and Vincke

[1] `http://www.fluka.org/flair/`, 27 November 2015.

2012] was developed at CERN in order to rapidly assess radiological hazards without the use of full Monte Carlo simulation codes. In the present thesis, we extensively used Actiwiz to predict radionuclide inventories and to estimate correlation factors. The detail of the calculations will be presented in the following sections and chapters.

2.1.3 Actiwiz

Actiwiz is a software developed at CERN to build a radiological hazard assessment for an arbitrary material exposed to the radiological environments of the accelerator complex [Theis and Vincke 2012].

The application was developed to give quick answers to general questions about radiological hazards without the need for the user to implement complex input files with a Monte Carlo code such as Fluka [Battistoni et al. 2006][Ferrari et al. 2005]. The developers have run thousands simulations representing 42 typical hadronic spectra (see for example Sec. 1.2.3) for various positions inside the accelerators' tunnels. The results of these simulations are stored as a database in Actiwiz and the user can run calculations on predefined simulated scenarios.

Among the available hazard factors, the user can calculate the IRAS and the fraction of LE (see Ch. 1, Sec. 1.3.2). To run a calculation, the user must enter the following information:

- machine and position (radiological environment);
- the material composition (mass weight of each element and density);
- the irradiation time t_i;
- the decay time t_c;
- the beam intensity (in lost particles per second).

The radiological environments available for the calculations performed for this thesis (version 2 of the code) represent all the accelerators of the CERN complex and include the Linac4 (160 MeV), the PS Booster (1.4 GeV), the PS (14 GeV\c), the SPS (450 GeV\c) and

the LHC (7 TeV). The positions considered are located on the beam impact area, within bulk material (e.g. magnet) surrounding the beam impact area, adjacent to bulk material surrounding the beam impact area, close to the concrete tunnel wall (loss on bulk object), behind concrete massive shielding, 10 cm lateral distance to a target, close to concrete tunnel wall (loss on target). Together, these radiological environments represent the set of possible realizations of the parameters E and P in the random vector scenario **S** of Eq. 2.1. These are discrete variables with respectively 5 and 7 possible outcomes.

Actiwiz calculations give a number of information which is valuable for the present thesis. For instance we can cite the radionuclide inventory, the contribution of each radionuclide to the IRAS and the list of target elements generating a given radionuclide.

The quantity "contribution to the IRAS" (C_i^{IRAS}) is relevant for the present thesis and we will often refer to it in the following chapters. It is defined as follows:

$$C_i^{IRAS} = \frac{\frac{a_i}{AL_i}}{\sum_j \frac{a_j}{AL_j}}. \tag{2.2}$$

where a_i is the activity of radionuclide i, AL_i is the coefficient expressing the radio-toxicity of radionuclide i (we will often refer to it as "activity limit") and the denominator is the $IRAS$ as given in Eq. 1.11. In practice, the contribution to the IRAS allows us to identify the radionuclides that have a real impact on the IRAS and to concentrate the characterization effort towards those radionuclides that have the strongest influences on the final hazard factor. As explained later, Eq. 2.2 is also used to discriminate between DTM and ITM radionuclides.

Before the analysis of the parameters of the random vector scenario, we present a practical example of the results that one can obtain from the code Actiwiz.

Example 1 *In this example we consider copper CuOFE or oxygen free, high-purity copper for electronic applications. The material composition is taken from a material catalogue commonly used at CERN [Froeschl et al. 2012]. This catalogue is employed as a reference for specific materials installed at CERN and it groups either producer's data or reference values from European or international standards on material compositions and requirements.*

Copper CuOFE consists of pure copper in which various trace elements are present. The trace elements (in weight fractions) are bismuth (0.1%), cadmium (0.01%), lead (0.1%), mercury (0.01%), oxygen (0.05%), sulfur (0.18%) and zinc (0.01%).

Using Actiwiz, we exposed copper CuOFE to the radiological environment of the LHC. The activation occurs behind the massive concrete shielding (tunnel wall), the material is irradiated for 7 days and it is left to decay for 2 years after the end of the irradiation. The list of the major elements produced, with their contribution to the total activity, is reported in Tab. 2.1.

Table 2.1: Radionuclide inventory generated by irradiation of copper CuOFE [Froeschl et al. 2012] behind concrete shielding of the LHC. The irradiation lasted for 7 days followed by 2 years of decay.

Radionuclide	**Contribution to total activity (in %)**
Co-57	36.89
Fe-55	20.63
Co-60	17.04
Mn-54	11.41
H-3	7.84
Ni-63	2.48
V-49	2.05
Co-58	1.04
Zn-65	0.33
Co-56	0.25
Ca-45	0.03
Sc-46	0.01
Total	100

Within the radionuclide inventory one can find ETMs such as Mn-54 and Co-60 and DTMs such as H-3, Fe-55 and Ni-63.

A second information given by Actiwiz is the contribution to the IRAS. This result is showed in Tab. 2.2. In this table one can see that Co-60 contributes above 50% to the IRAS, even if Co-57 is the major contributor in terms of activity (see Tab. 2.1). As it is discussed in Sec. 1.3.1, each radionuclide has a different activity limit AL that is used to calculate the IRAS. For instance, Co-57 has an AL equal to 100 and the Co-60's AL is 10. Because of its activity limit Co-60 contributes more to the hazard factor. Radionuclides such as H-3, Fe-55 or Ni-63 contribute even less to the IRAS because their activity limit is 1000.

Another information that we can obtain from Actiwiz is the contribution of each single

Table 2.2: IRAS top contributors obtained by activating copper CuOFE behind concrete shielding of the LHC. The irradiation lasted for 7 days followed by 2 years of decay.

Radionuclide	$\mathbf{C_i^{IRAS}}$ **(in %)**
Co-60	52
Mn-54	35
Co-57	11
Total	98

chemical element in the production of a given radionuclide. In our example, Actiwiz shows that Co-60, Mn-54 and Co-57 are produced almost exclusively by the element copper (∼100%). Small contributions come also from zinc (∼ 10^{-5}*%) and lead (∼* 10^{-7} *%).*

□

Actiwiz was initially created as a graphical user interface (GUI) based application. However, when performing large amount of calculations, involving dozens chemical compositions and thousands of different scenarios, the Actiwiz GUI cannot be used efficiently. For this reason, internal scripts were developed by the authors of the code to access its core functionality. These scripts, written using the programming language Python[2] [Rossum and Drake 2001], have the advantage of generating simultaneously the desired outputs for as many scenarios as wanted. In the frame of the present thesis we developed advanced scripts for post-processing of data, including complete statistical analysis using the code R[3] [R CORE TEAM 2014]. These tools are nowadays routinely used at CERN for waste characterization.

2.2 Elemental composition

The knowledge of the elemental composition of an item is crucial to study the production of radionuclides when the material is exposed to energetic beam particles, such as the ones produced at CERN. The exact elemental composition of legacy waste is however very difficult to estimate for various reasons. First of all, with the exclusion of the major elements, it is often either technically or economically cumbersome to estimate trace

[2] `https://www.python.org/`, 27 November 2015.
[3] `https://www.r-project.org/`, 27 November 2015.

elements on large amounts of materials. Moreover, the quantity of some trace elements is well below the detection limits of common instruments (for instance, the detection limit of Co in steel, for common portable instruments employed at CERN, is above 50 ppm). On the other hand, the variability of trace elements can be very high and it can be hard to collect a representative sample from a waste population, especially if limited process knowledge is available.

When building a chemical composition for activation studies it is important to include all the known elements but also the traces and the suspected contaminants. It is however possible to screen out some of the elements that either do not produce radionuclides of interest for the characterization process or if the radionuclides produced have a very small impact on the hazard factor.

The collection of elemental composition data can be made via direct measurement, collection of data from literature and producers or from original material specifications and national or international standards. The data can be either in the format of a single value, like the average value, or of a distribution or range.

Traditionally, the studies carried out at CERN for waste characterization used "standard" chemical compositions from a material catalogue [Froeschl et al. 2012]. This catalogue collects producer's data and values from European and international standards for materials that are used to build accelerator's components and structures. This catalogue gives, for each material, an average value of percentage weight per element. The catalogue is extensive but, by definition, gives only average common values for each element and for a specified composition.

A second approach to identify elemental compositions consists of building a distribution for each element, as outlined in Tab. 2.3. This approach is described in the standard of reference [ISO 2013].

In practice we will adopt one method or the other depending on the amount of information or data (such as measurements) we have on the elemental composition of the waste under study. According to reference [ISO 2013] we can identify the 4 following cases.

- Element analysis data is sufficient and representative. In this case it is possible to set

Table 2.3: Philosophy for setting basic type of distributions of chemical elements. Adapted from [ISO 2013].

Chemical element condition	**Main elements. Controlled in a certain range of concentration**	**Impurity elements. Controlled with an upper limit of concentration**	**Trace elements. Non-controlled**
Basic philosophy	Main chemical elements of materials which are manufactured in specific factories under lot-based quality control. Their contents are controlled within the target ranges specified by national industrial standards of material, and their concentration ranges are comparatively narrow.	Chemical elements which are reduced or controlled in a certain manufacturing process as impurity elements contained in manufactured materials. Their contents are controlled below comparatively low control values, and the concentration distribution of each element is able to reflect its concentration distribution in the nature.	Chemical elements not controlled. The content of each element reflects its concentration distribution appearing in nature.
Reference concentration distribution of each chemical element	Normal distribution	Log-normal distribution	Log-normal distribution

the distributions using average values, standard deviations, ranges or minimum/maximum values for each element.

- Element analysis data is limited. In this case conservativeness can be used by setting representative values as being the upper limits of the confidence intervals obtained from data for each element.
- Data is very limited (very few). When the number of measured values is very limited, it is possible to either use the detected values to estimate averages and standard deviations of the distributions, or to assume that the concentration distribution lies in a range below the detection limit. For instance, it is possible to make the hypothesis that the maximum detected value lies at 2σ from the centre of the distribution and the standard deviation can be evaluated from the same element data in nature or from element data which are chemically similar.
- Element analysis data contains detection limit values only. In this case we can either use the detection limits to estimate averages and standard deviations or we can generate distributions that lie below the detection limit (similarly to the case when data are very few) or we can perform radiochemical analysis on target radioactive waste whose irradiation condition is known.

As previously stated, it is often difficult to obtain information about the exact chemical nature of legacy waste. This is mainly due to the absence of a traceability system in place at the time of the waste generation. Even if the data is available, it is often partial and the link between the data and the material stored is difficult to establish. For these reasons, we are commonly confronted with inference based upon limited amount of measurements which are performed on accessible materials (with all the limitations of biased sampling).

At this stage it is crucial to consider an extended list of potential elements that can have an impact on the radiological characterization of the waste. To find elemental compositions of metals, it is possible to use both standards and analyses performed at mining and industrial sites where the metals are extracted and refined. It is also important to understand the production mechanisms of the metals because they allow us to justify the presence of trace elements but also contaminants.

We show in the following detailed example how the techniques previously described can be applied to generate the distributions of the elements in cathodic copper, which is commonly used at CERN for electric and electronic applications.

Even if the approach that we present is specific for cathodic copper, the steps of the procedure are similar when building other elemental compositions for activation studies.

Example 2 *Copper is extensively used at CERN in a number of applications spanning from cables to machine's components. In this example we consider high-purity copper because of its relevance on radioactive waste characterization. We show in particular how a random chemical composition of copper can be built using both experimental data and values from producers and standards.*

Copper is present in the earth's crust as copper-iron-sulfide and copper sulfide minerals, e.g. chalcopyrite ($CuFeS_2$*), bornite (*Cu_5FeS_4*) and chalcocite (*Cu_2S*). The concentration of these minerals in an ore body is low. Typical copper ores contain from 0.5% Cu (open pit mines) to 1 or 2% Cu (underground mines). Pure copper metal is produced from these ores by concentration, smelting and refining [Schlesinger et al. 2011].*

Copper also occurs in oxidized minerals like azurite ($Cu_3(CO_3)_2(OH)_2$*), but to a lesser extent. Copper metal is usually produced from these minerals by hydrometallurgical methods.*

A third major source of copper is scrap copper and copper alloys. Production of copper from recycled used objects is 10 or 15% of mine production. In addition, there is considerable re-melting/re-refining of scrap generated during fabrication and manufacture [Schlesinger et al. 2011].

The purity of copper is very high and it is commonly above 99.9% [ASTM 2016], independently from the production method used. The list of trace elements and impurities can be however very long and can vary from extraction site to extraction site and depending on the production method and the refining technique used. A complete list of impurities in cathodic copper can be found in various European and American standards (such as CEN-EN 13347 [CEN 2002] and ASTM B115-10(2016) [ASTM 2016]). For the interested reader, references [Schlesinger et al. 2011][CDA 2004] give a precise and extensive discussion about amount and origin of each trace element. In the present thesis the list

of impurities considered includes the elements Ag, As, Bi, Cd, Co, Fe, Ni, O, Pb, S, Sb, Se, Sn, Te and Zn. These impurities are systematically below the detection limits of the instruments used at CERN (above 50 to 100 ppm for most elements).

We can identify three types of elements, when considering the information collected about the chemical composition of legacy waste made of copper. Each type or family is discussed individually.

Measurable principal chemical elements.

The main element copper is easily measurable and its mean and standard deviation are obtained from an incremental collection of data on random samples. For this element we can build a normal distribution $\mathcal{N}(\mu, \sigma)$ *(see Fig. 2.4) according to the indication of Tab. 2.3. It is easy to see that the experimental average concentration of copper (99.88%) is very close to the expected values (99.9%) given by the reference standards [ASTM 2016] [CEN 2002]. The standard deviation of the measurements (0.11%) is given for a cover factor* $k = 1$.

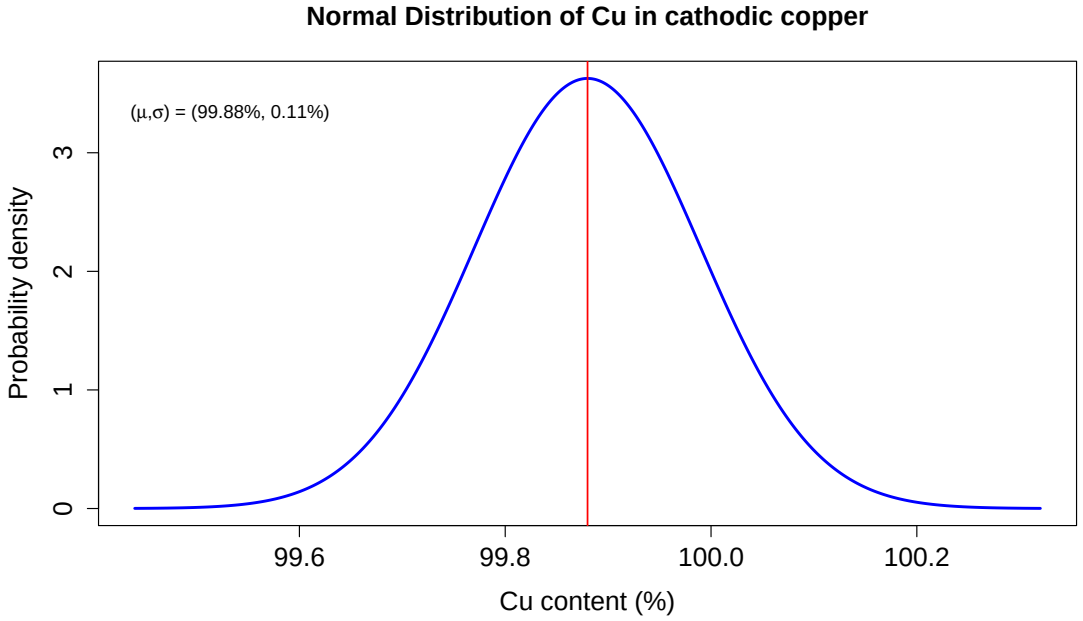


Figure 2.4: Experimental normal distribution of cathodic and unalloyed copper from measurements on incremental random samples performed on radioactive cables.

When copper alloy materials such as Cu-Al, Cu-Ni, Cu-Sn and Cu-Zn are considered, the alloying elements must be quantified using a similar approach to that used for copper. Normal distributions can be easily built using experimental data from measurements as it is showed for the element copper.

Elements below detection limits for which data about distribution exists.

It is common to evaluate the amount of trace elements of metals via measurements at production and mining sites. Impurity control is essential because the quality of a material is correlated to the number and the concentration of minor elements. Some elements, like Ag and Au in copper, are relatively benign, because they do not affect the electrical/electronic properties of the metal. The recovery of precious metal is however often important for their intrinsic value. On the other hand the presence of impurities like Bi, Pb and Sb induces severe brittleness and Se and O affect copper ductility [Camurri et al. 2011]. Detailed concentration data for various trace elements and impurities can therefore be found in literature and can be used to build experimentally based distributions. The data can give either the element concentration of the ores (natural distributions) or their amount within the concentrates. The use of data from concentrates reflects a conservative approach because the amount of impurities and trace elements are not yet minimized via either smelting and refining or hydrometallurgy.

Tab. 2.4 shows the elemental composition of world-wide copper concentrates collected by the European Copper Institute [CDA 2004]. The amount of Cu in the concentrates considered is in the range 14-51% with median value 26.7%.

Table 2.4: Elemental composition of world-wide copper concentrates [CDA 2004]. The values of the quantiles for a given percentile p are in %. The number of samples considered for the analysis is $n = 119$.

n=119	**Sb**	**As**	**Zn**	**Pb**	**Ni**	**Ag**	**Cd**	**Co**
min	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
p50%	0.010	0.110	0.620	0.140	0.002	0.006	0.004	0.005
p60%	0.015	0.139	1.307	0.266	0.004	0.008	0.006	0.009
p70%	0.022	0.18	2.872	0.562	0.008	0.011	0.010	0.013
p80%	0.042	0.272	3.652	1.478	0.010	0.017	0.014	0.024
p90%	0.102	0.410	5.632	2.910	0.024	0.068	0.026	0.040
max	7.250	7.5	9.280	12.710	1.030	1.907	0.072	0.250

Various algorithms exist to determine the parameters of a location-scale family distribution from quantiles. In the present study we used the R-language [R CORE TEAM 2014] implementation of the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm, which is an iterative method used in numerical optimization for solving unconstrained nonlinear optimization problems [Avriel 2003] [Broyden 1970] [Fletcher 1970]. If the BFGS algorithm fails to converge, the Nelder-Mead method is used instead [Nelder and Mead 1965].

According to reference [ISO 2013] (see Tab. 2.3) we are interested on log-normal distributions for the trace elements. If we indicate with μ the mean and with σ the standard deviation of the variable's logarithm, we can write the major statistics of a log-normal distribution as follows:

$$E[X] = e^{\mu + \frac{\sigma^2}{2}} \tag{2.3}$$

$$Med[X] = e^{\mu} \tag{2.4}$$

$$Mode[X] = e^{\mu - \sigma^2} \tag{2.5}$$

$$Var[X] = e^{(2\mu + \sigma^2)} \times (e^{\sigma^2} - 1) \tag{2.6}$$

$$SD[X] = \sqrt{Var[X]}. \tag{2.7}$$

Tab. 2.5 gives the parameters of the log-normal distributions obtained for the elements showed in Tab. 2.4. For the calculations we replace the min and the max with the percentiles p1% and p99% respectively. The values corresponding to the min where also replaced by 10^{-8}% for numerical reasons. The probability density plots obtained are showed in Fig. 2.5.

We can test the accuracy of these distributions comparing their median (Eq. 2.4) or the mode (Eq. 2.5) with the upper limits of impurities in cathodic copper set by international standards. We do not compare the expectation with the limits because this central tendency

Table 2.5: Distribution parameters (μ, σ) obtained from quantiles for impurities in world-wide copper concentrates.

Parameters	**Sb**	**As**	**Zn**	**Pb**	**Ni**	**Ag**	**Cd**	**Co**
μ	-4.65	-4.11	-2.23	-1.93	-6.12	-5.21	-5.49	-5.23
σ	1.75	3.69	4.15	2.48	2.02	1.51	1.48	1.71

estimators is strongly affected by extreme values in right-skewed distributions, such as log-normal distributions. The comparison of distribution parameters and values from standards is given in Tab. 2.6.

Table 2.6: Comparison between median and mode of the distributions of trace elements obtained from copper concentrates [CDA 2004] and maximum allowed impurities in electrolytic cathodic copper (Grade A) as specified by the standards ASTM B115 [ASTM 2016] (Grade 1 and 2) and [BS 1998] (Cu-CATH-1). The values are in ppm.

El.	**Med[X] (ppm)**	**Mode[X] (ppm)**	**Gr. 1 (ppm)**	**Gr. 2 (ppm)**	**Cu-CATH-1 (ppm)**
Ag	55	6	25	70	25
As	165	2×10^{-4}	5	15	5
Bi	-	-	1	3	2
Cd	41	5	-	-	-
Co	53	3	-	-	-
Fe	-	-	10	25	10
Ni	22	0.4	10	20	-
O	-	-	-	-	400
Pb	1455	3	5	40	5
S	-	-	15	25	15
Sb	96	5	4	15	4
Se	-	-	2	10	2
Sn	-	-	5	10	-
Te	-	-	2	5	2
Zn	1071	3×10^{-5}	-	-	-

In Tab. 2.6 we compare the distribution statistics with three grades from the reference standards of electrolytic cathodic copper. Grade 1 is a high purity copper that is used to manufacture wire rods. Grade 2 has similar characteristics of grade 1 but it is used for less demanding applications. Cu-CATH-1 is reported because it gives the numeric value of maximum oxygen content.

From the table one can see that the calculated mode (maximum of log-normal distributions) is below the maximum allowed impurities as given by the reference standards with the only exception of Sb. Also in this case however the values are very close.

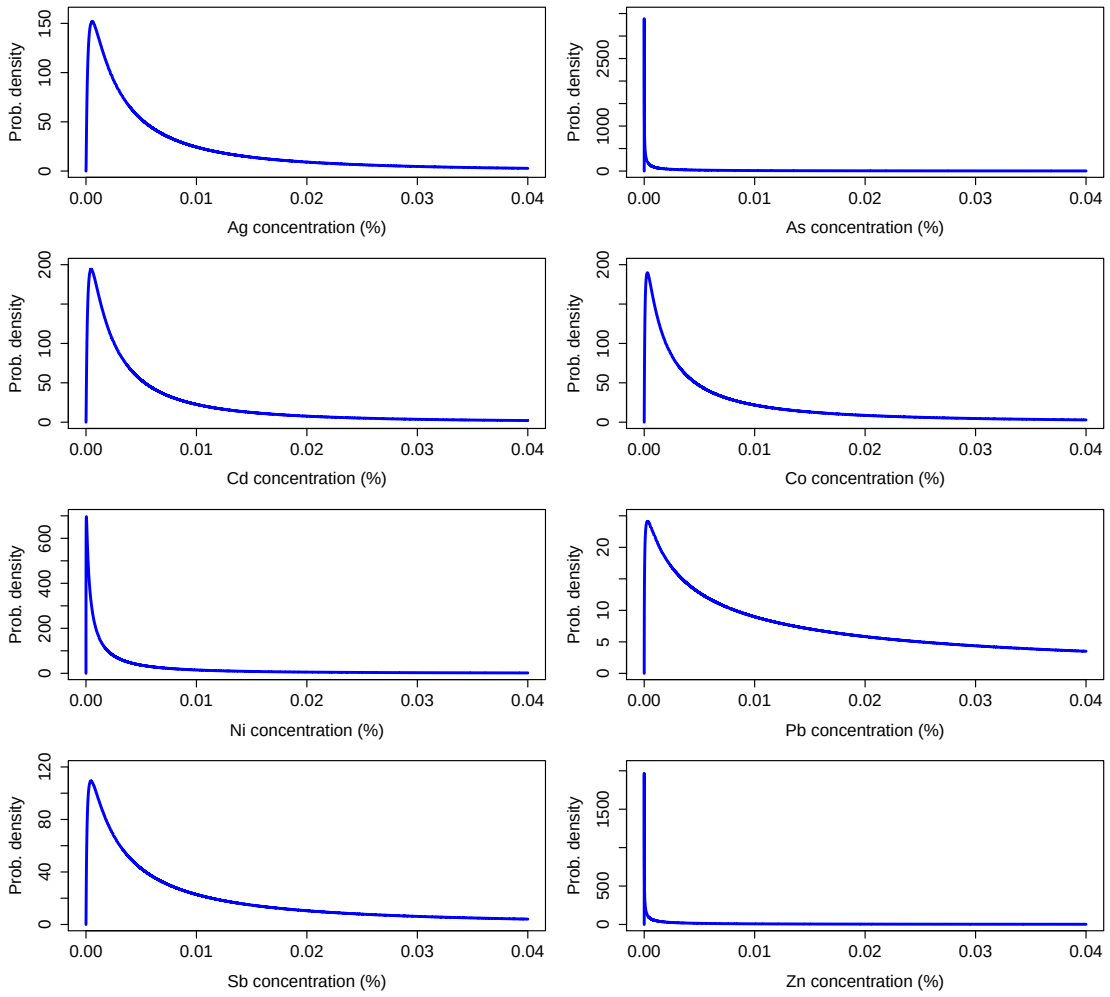


Figure 2.5: Calculated distributions of impurities from concentrates of copper [CDA 2004]. The probability densities are built from 119 representative samples.

Experimental data from concentrates is in very good agreement with the maximum allowed values given by the norms and is a good indicator of the robustness of the methodology used to build the distributions. Some impurities for which we do not have a value in the standards (Cd, Co and Zn) are present on the data from concentrates. This is due to the fact that, at this stage, the concentrates need to undergo purification processes before being available with the purity levels required on the final product.

Elements below detection limit for which data about natural distribution is not available.

A last case occurs when experimental data for trace elements and impurities is not available. For these elements we can consider the upper concentration limit fixed by the reference standards. These limits are respected by the producers because the properties of the materials (mechanical and electric/electronic) can drastically change as previously discussed. In this case a single value can be used as representative for the trace element.

We will use this approach for the elements Bi, Fe, O, S, Se, Sn and Te in cathodic copper.

□

In the previous example we showed how we can estimate either the distribution or the value of each chemical element in a given material composition. In practice, when we have identified a waste population (step A of the characterization process), we perform a limited number of measurements to identify the major chemical elements. For minor elements which are present as a trace or an impurity we can either collect data from literature to build a distribution or we can use standards for average values. At the end of this process we have compositional values as either single values or distributions. These data can be used to build elemental compositions with stochastic variations.

Let us consider a generic elemental composition with n elements EL_i $(i = 1, 2, ..., n)$. EL_i represents the vector of realizations for the i^{th} element. Suppose also that we want to generate N random elemental compositions that we will identify with **CC**. If we indicate with $\mathcal{P}(\mu_i, \sigma_i)$ the probability density function from which a realization is withdrawn we can write:

$$EL_i = (x_{i1}, x_{i2}, \cdots, x_{iN}) \overset{r}{\leftarrow} \mathcal{P}(\mu_i, \sigma_i). \tag{2.8}$$

Eq. 2.8 indicates that the N realizations of the vector EL_i are obtained from a random extraction ($\overset{r}{\leftarrow}$) of values from the probability density function $\mathcal{P}(\mu_i, \sigma_i)$. We have seen that $\mathcal{P}(\mu_i, \sigma_i)$ is either a normal distribution (major measurable elements) or log-normal (impurities for which experimental data of natural compositions is available). When only the maximum or the average amount M_j of the element j is available, the probability

density function is replaced by a vector in which M_j is repeated N times.

Such a matrix has the capability to include stochastic variations of the elemental composition in the input of the activation studies. A second advantage of such a method is that elemental composition random vectors reflect more accurately the real heterogeneous nature of materials. Finally, however improbable, realizations with high content of trace elements can be randomly extracted and used to predict particularly disadvantageous radionuclide inventories that can be seen as extreme boundary cases.

As an example, we built a random vector for the chemical composition of cathodic copper with $N = 100$. The results are presented as a matrix in which the row i gives the N random realizations of element i and the column j indicates a generic random chemical composition. Each column is normalized to add up to 100%.

$$\begin{array}{c} \\ Ag \\ As \\ \vdots \\ Cu \\ \vdots \\ Zn \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & \cdots & \mathrm{N} \end{array} \\ \begin{pmatrix} 0.10613 & 0.00532 & \cdots & 0.00668 \\ 0.00155 & 0.00058 & \cdots & 0.89897 \\ \vdots & \vdots & \cdots & \vdots \\ 98.82159 & 99.60654 & \cdots & 98.58783 \\ \vdots & \vdots & \cdots & \vdots \\ 0.43845 & 0.00998 & \cdots & 0.03522 \end{pmatrix} \end{array}$$

This matrix represents the mixed multivariate random chemical composition that can be used for analytical calculations or Monte Carlo simulations when predicting the nuclide database of materials activated at CERN.

The elemental composition of a radioactive waste is only one of the input parameters needed for the calculations. In the following sections we will describe the energy, location, irradiation and decay times that, together with the elemental composition, are necessary for the calculation of the radionuclide inventory, the correlation and the scaling factors.

2.3 Position within the tunnel and beam energy

The energy of the accelerator and the position of the material within the tunnel play an important role on the activation mechanisms. This is due to the type, number and quality

of secondary particles that can be generated. In Sec. 1.2.3 we showed two examples of particle spectra generated at two distinct locations within the LHC. As already discussed, even if it is possible to generate particle spectra for each single position and each accelerator using Monte Carlo codes, this would be cumbersome without really adding a valuable information on the radiological characterization of legacy waste. In order to study the activation of materials within the tunnels, it is however essential to identify locations for which the spectra and the activation mechanisms are fairly representative of the reality.

This simplification task was carried out by the developers of the analytical code Actiwiz [Theis and Vincke 2012], which is based on intensive Monte Carlo simulations and described in Sec. 2.1.3. Fig. 2.6 presents the 7 positions within the accelerator tunnel as implemented in Actiwiz, version 2 [Theis and Vincke 2012]. These 7 positions, together with the energies of the accelerators are the possible realizations of the input parameters identified respectively with P and E in Eq. 2.1.

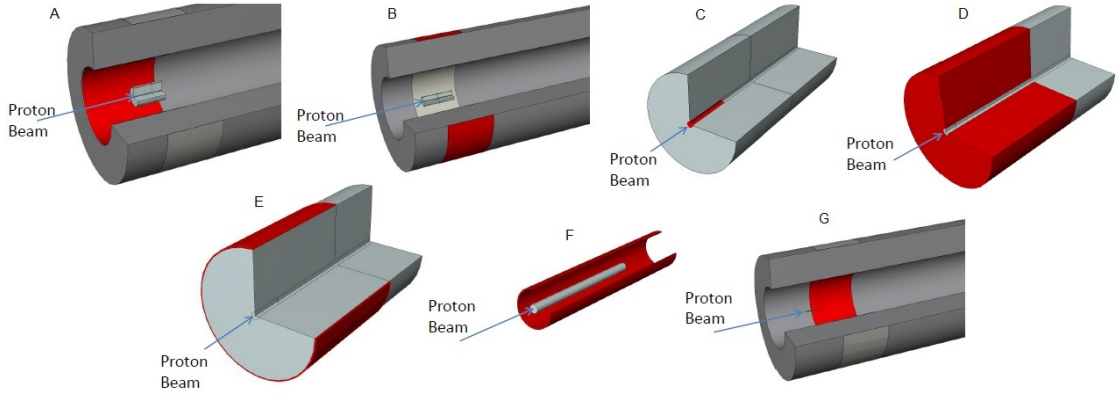


Figure 2.6: Standard locations implemented in Actiwiz (version 2) [Theis and Vincke 2012] to score particle fluences within the CERN's accelerator complex.

The reference positions showed in Fig. 2.6 are simplifications of specific locations of the accelerator complex. The list that follows (letters A to G) gives the details of each one of them. The volume used for fluence scoring is pictured in red.

A. The thin red hollow cylinder is located close to the tunnel wall at a lateral distance to the beam axis of 2 m, starting 2 m upstream the beam impact point (length of 4 m). This irradiation situation can be used to describe materials located at the tunnel

wall when the beam energy is lost in massive objects like magnets. The scheme is specific for primary protons of 1.4 GeV accelerated at the PS Booster.

B. The thin red hollow cylinder is located behind 2 m of lateral shielding (length of 3 m). This irradiation situation can be used to describe materials located behind thick lateral concrete walls which shield radiation from beam impacts. The scheme is specific for primary protons of 400 GeV/c accelerated at the SPS or CNGS.

C. The red cylinder is located at the beam impact point. This irradiation situation can be used to describe materials suffering direct beam impact. The scheme is specific for primary protons of 14 Gev/c accelerated at the PS.

D. The red hollow cylinder is located close to objects intercepting the beam (3 cm $<$ r $<$ 50 cm). This irradiation situation can be used to describe solid massive materials located close to objects intercepting protons from the beam line. The scheme is specific for primary protons of 160 MeV accelerated at the Linac 4.

E. The red hollow cylinder is located beside massive equipment like magnets (50 cm $<$ r $<$ 51 cm). The scheme is specific for primary protons of 1.4 GeV accelerated at the PS Booster.

F. The red hollow cylinder (thickness of 1 cm) surrounds the target. This irradiation situation can be used to describe materials located close to objects with small lateral extension like targets or collimators. The scheme is specific for primary protons of 7 TeV accelerated at the LHC.

G. The thin red hollow cylinder is located close to the tunnel wall at a lateral distance of 2 m from the beam axis and starting 1 m downstream to the beam impact. This irradiation situation can be used to describe materials located at the tunnel wall in areas where beam is lost in objects like targets and collimators. The scheme is specific for primary protons of 7 TeV accelerated at the LHC.

When a radioactive waste is received at the CERN temporary storage, its origin is registered together with other useful information like the average and maximum equivalent

dose rate measured (at contact of the waste and at 10 cm) the day of the reception. For this reason, the parameter E is often known even though, for old legacy waste, the reliability of this information cannot be verified. The position P of legacy materials within the tunnel, on the other hand, is often missing and cannot be established with certainty when performing calculations.

For the reasons mentioned, when withdrawing random values for E and P we simply attribute the same probability to each potential outcome. In particular, we have 7 outcomes for the position and 5 for the energy. We would like to point out a few limitations of this method.

First of all, the regions for fluence scoring are not real but simplified regions which can be thought as representative for common CERN installations. Materials however can be activated in areas that are intermediate between two or more of the regions considered with an impact on the correspondent particle spectrum. We should also cite that the real position of the object under consideration could be upstream or downstream of the beam loss.

An equal probability between locations represents a bias because an equal amount of materials would be activated independently from their distance from the beam line. This is not true in the sense that only a limited amount of materials are directly exposed to the beam while the majority of activated objects lie at a greater distance from the beam line. To overcome this problem it can be possible to weight each potential outcome of the position P according to the volume occupied by the object on the 7 reference positions. This approach is however not yet implemented.

For completeness we should add that the definition of the 7 standard positions is a feature of Actiwiz version 2. The future version 3 will allow the user to define as much positions as desired because it will be able to accept generic particle fluence spectra, preliminarily calculated with Fluka.

Finally, a few limitations occur also with the implementation of real scenarios when using the energies available within Actiwiz. This is due to the absence of fluence particle spectra for Linac 2, the first stage of acceleration of the CERN accelerating complex. As

previously discussed in Ch. 1, Linac 2 will be replaced by the new Linac 4 between 2017 and 2018. For this reason the authors of Actiwiz decided to implement the analytical tool considering only the new linear accelerator. Differences in particle spectra are therefore to be expected between these two accelerators.

2.4 Irradiation and decay times

The last two parameters required in Actiwiz to calculate the radionuclide inventory are irradiation and decay times.

The irradiation time t_i is defined as the time during which an item or an object is directly or indirectly exposed to a radiation field. It is very difficult to evaluate the irradiation time of a single object removed from the tunnel because it depends on the so-called irradiation cycle of the accelerator, on when it was installed, on when it was removed and on the time it spent outside a radiation area for maintenance. Accelerators work continuously over the years except when they undergo maintenance (small and long shut-downs). The time in which items are not irradiated because of regular maintenance is however very short in comparison to the actual irradiation time. For this reason it can be assumed that the irradiation of the majority of objects is a continuous phenomenon with the time. For legacy waste the irradiation time is often unknown because these parameters had not been recorded.

For the present study we consider that the irradiation time spans from 1 second up to 40 years. The upper limit is set to cover some structural elements which can have a lifespan of 40 years. However, machine elements are commonly replaced during maintenance or upgrades and can rarely reach such a high value of irradiation time. For operational calculations the upper limit is set to 30 years.

The decay time t_c is a crucial parameter for the characterization of radioactive waste. It is defined as the time elapsed since the end of irradiation.

As for the irradiation time, the decay time is difficult to estimate for legacy waste. A lower limit can be set by using the date when an object is received at the temporary storage.

In Fig. 2.7 we can see the histogram of items recorded as waste after reception at the CERN's storage centre. The number of items received since the foundation of the laboratory amounts to over 15500 units.

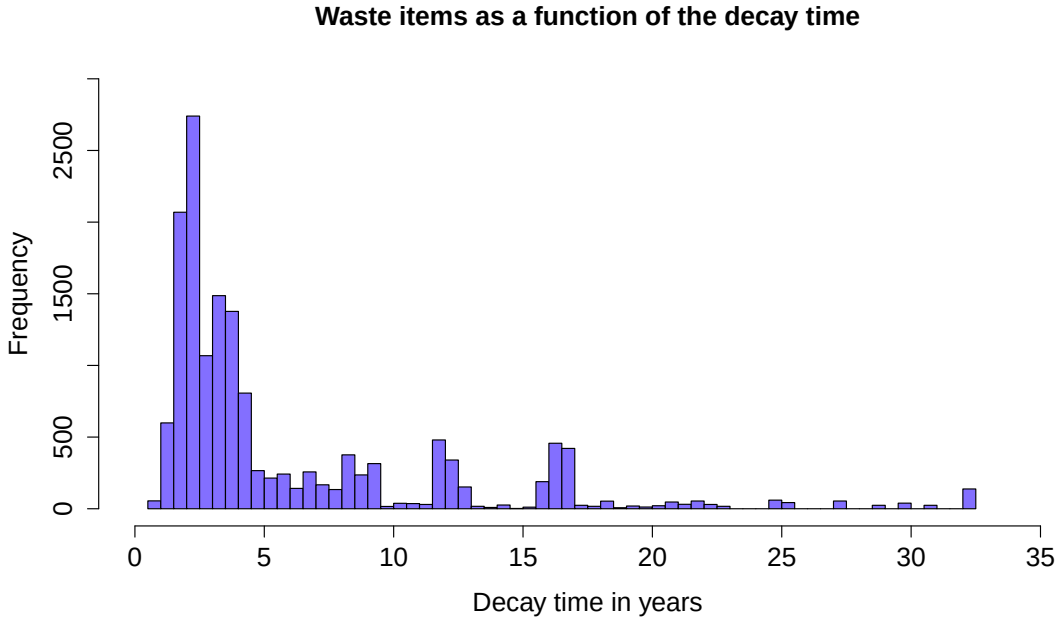


Figure 2.7: Histograms of waste items recorded as a function of the decay time estimated from the date of reception at the CERN's storage centre.

The volume of each item can span from a few cubic centimetres up to dozen cubic meters. The plot shows a major peak between 2-to-3 years decay time (2013-2014) and various small peaks around 12 and 16 years decay time (years 2000/2001 and 2005). Regular inventory campaigns were started over the years generating the peaks that we can see in figure. For instance, between 2012 and 2014 the first long shut-down to increase the energy of the LHC produced more than 2500 tons of waste. A new system to trace materials (called TREC) was also put in place during those years. The combination of the new traceability system together with the technical shut-down accounts for the biggest peak in Fig. 2.7. Similarly we can explain the peaks at the years 2000/2001 (dismantling campaign of the LEP) and 2005 (inventory of waste operated by the National Cooperative for the Disposal of Radioactive Waste, NAGRA, CH[4]).

[4] `http://www.nagra.ch/en`, 6 February 2017.

For legacy waste it is difficult to set a precise decay time. The graphic of Fig. 2.7 can however be used to identify the span of the decay time which is set within the range 0-40 years for general calculations and 0-30 years for operational calculations. For both irradiation and decay times, in absence of better data, we consider the distribution uniform in the entire range.

2.5 Selecting the number of realisations

At this stage of the discussion we have introduced the characteristics of each parameter of the random vector **S**, the typical distributions from which we can withdraw random realizations and the reasons that made us set boundaries on some parameters such as the irradiation and decay times.

The range methods, introduced at the beginning of the present chapter, can be used to estimate the activities of the radionuclides in a waste when we are confronted with random input parameters. The basic idea of such an approach is the random extraction of input variables from the input space to collect an appropriate set of input conditions. The input conditions, consisting of a given number of realizations of the scenario **S**, can be used as input parameters for either Monte Carlo or analytical calculations.

The outcome of the simulations or calculations are distributions of specific statistics such as the average correlation factor between KN and DTM radionuclides. In this section we discuss the criterion adopted to select the number of random extractions to build distributions.

The correlation method [ISO 2013] indicates that, under certain conditions, it is possible to establish a correlation between the activities of a key nuclide and difficult-to-measure radionuclides based on theoretical activation studies. For a fixed realization i of the random vector scenario we can calculate the correlation factor CF_i as follows:

$$CF_i = \frac{a_{DTM_i}}{a_{KN_i}} \tag{2.9}$$

where a_{DTM_i} is the activity of a generic DTM radionuclide for the theoretical scenario i and a_{KN_i} is the calculated activity of the KN for the same scenario.

If we extract a large number of realizations of the random vector scenario we can build a distribution for the CF. On over 2.35 million simulations performed we found that these distributions are often right skewed and follow a log-normal distribution. For log-normal distributions, a robust central tendency estimator is the so-called geometric mean.

For log-normally distributed CFs we can calculate the geometric mean as follows [ISO 2007] [ISO 2013] [IAEA 2009b]:

$$\overline{G}_{CF} = exp\left(\frac{1}{n}\sum_{i=1}^{n}\ln(CF_i)\right) = \left(\prod_{i=1}^{n} CF_i\right)^{\frac{1}{n}}. \tag{2.10}$$

The equality of Eq. 2.10 is valid only if each $CF_i > 0$.

In order to fix the number of scenario's realizations we can check the variation of the parameters of the log-normal distribution while increasing the number of random extractions. We can finally fix this number when the parameters stabilize below a fixed level of variation. For log-normal distributions we are interested on the stabilization of the geometric mean, given in Eq. 2.10, and of the geometric standard deviation, also called dispersion D. For correlations factors, the dispersion can be written as follows:

$$\mathrm{D} = exp\left(\sqrt{\frac{\sum_{i=1}^{n}[\ln(CF_i) - \ln(\overline{CF})]^2}{n-1}}\right) \tag{2.11}$$

where $\overline{CF}$ is the average correlation factor calculated over the n realizations.

We illustrate the strategy for choosing the number of realizations via an example.

Example 3 *We want to choose an adequate number of realizations when considering cathodic copper irradiated at CERN. Common DTMs produced by activation of cathodic copper are Ni-63 and H-3. A standard KN is Co-60. In particular, we are interested in the minimum number of scenarios that should be considered, when establishing correlation factors, to avoid high random variations on the the geometric mean correlation factor and its dispersion.*

The method that we adopted consists of calculating the geometric mean correlation factor and the geometric standard deviation multiple times. At each calculation we withdraw

an increased number of scenarios and we compare the results obtained with the ones from the previous stage.

In Fig. 2.8 we show the results obtained for the geometric CFs and the geometric standard deviations of Ni-63 and H-3 from cathodic copper. The parameters are calculated for a number of scenarios spanning from 10 up to 10000. The plots at the top of the figure show the evolution of the CFs for Ni-63 (top-left) and H-3 (top-right). The plots at the bottom of Fig. 2.8 (in red) show the evolution of the geometric standard deviations for Ni-63 (bottom-left) and H-3 (bottom-right) with the number of scenarios.

From the plots it can be seen that the parameters of the distributions stabilize roughly above 1000 scenarios.

□

A more precise estimation that the one proposed in the previous example can be made calculating the relative error ϵ_{rel} between two consecutive fixed number of realizations, and introducing a threshold l below which ϵ_{rel} must remain.

The relative error on the correlation factor between two consecutive fixed number of realizations x and y is defined as follows:

$$\epsilon_{rel} = \frac{|CF_y - CF_x|}{CF_y} \tag{2.12}$$

where CF_y and CF_x are the correlation factors calculated with y and x realizations (with $y > x$).

In a similar way we can define the relative error of consecutive geometric standard deviation. Finally, we fix the limit l as low as possible (generally below 5%) so that the calculations can be performed within a reasonable computer running time (less than 24 hours depending on the material considered and for one dedicated 3.2 GHz CPU).

Using Eq. 2.12 we found that more than 5000 scenarios should be considered when estimating CFs if we want to ensure a relative error below 5%. To be on the safe side, we fixed at 10000 the number of realizations to be considered when performing such a kind

of calculations. This number is also acceptable in terms of computing time and memory capacity of desk computers.

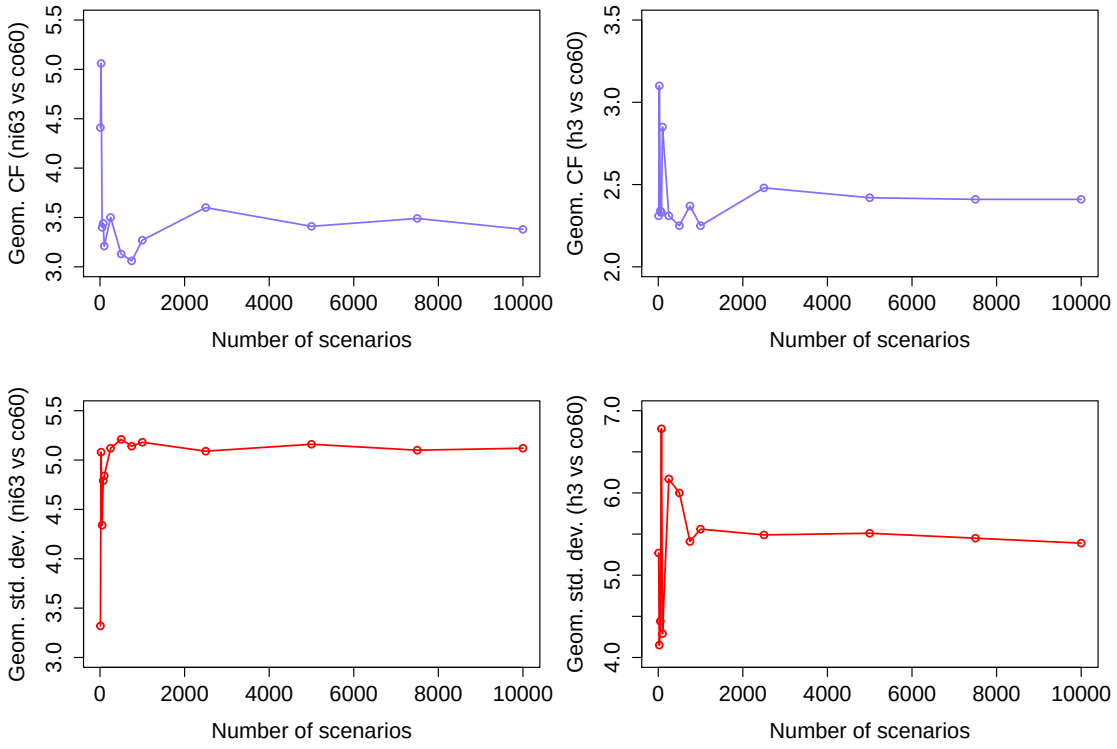


Figure 2.8: Selection of the number of scenarios using the stabilization of the parameters of the log-normal distribution. In blue (top) are given the geometric mean CFs for Ni-63 and H-3 in cathodic copper. In red (bottom) are presented the variations of the geometric standard deviations for the same radionuclides with the number of scenarios.

2.6 Radionuclide inventory of cathodic copper

We conclude this chapter with an example that summarizes the concepts introduced. We show the radionuclide inventory produced by activation of cathodic copper at CERN. For this example we consider 10000 random scenarios according to the rule described in Sec. 2.5.

The radionuclide inventory generated in a material exposed to a large number of possible CERN's activation scenarios can be long (it typically includes above 50 radionuclides). However, only a limited number of radionuclides have a real impact on the hazard factor

associated with the waste packages and have an activity above the Declaration Threshold. One way to quantify this impact is to estimate the contribution C_{IRAS} of each radionuclide to the IRAS as described in Sec. 2.1.3, Eq. 2.2.

Example 4 *Tab. 2.7 shows the list of radionuclides that contribute at least 0.05% to the IRAS when 10000 random activation scenarios are extracted. The model adopted to extract the input parameters is described in the list that follows.*

Table 2.7: Radionuclide inventory of cathodic copper obtained extracting 10000 random realizations from the random vector scenario. Only radionuclides contributing at least 0.05% to the IRAS are showed.

Radionuclide	**ETM**	**DTM**	**ITM**	**C_{IRAS} (in %)**
Co-60	✓			93.91
Ni-63		✓		3.17
H-3		✓		2.27
Ag-108m	✓			0.35
Ti-44	✓			0.20
Fe-55			✓	0.05
Total				99.95

- *The material composition (**CC**) is obtained by random extraction from a list of chemical elements. The main element copper is obtained from random extraction from the experimental normal distribution given in Sec. 2.2. The elements Ag, As, Cd, Co, Ni, Pb, Sb and Zn are obtained by random extraction from log-normal distributions built using reference data as discussed in Sec. 2.2. Finally, we attributed values from the reference standards to the trace elements for which experimental data is not available. In particular, we used the values of the standard ASTM B115-10(2016) [ASTM 2016] for Bi, Fe, O, S, Se, Sn and Te. It should however be outlined that at the time of writing (2017), the cross-sections of the elements Se and Te are not available for calculations in Fluka [Battistoni et al. 2006][Ferrari et al. 2005] and, by consequence, in Actiwiz [Theis and Vincke 2012].*

- *The energy (E) and position within the tunnel (P) are obtained as random extractions from the discrete vectors discussed in Sec. 2.3. Each potential outcome (7 for the position and 5 for the energy) is equally probable.*

- *Irradiation (t_i) and decay (t_c) times are obtained as random extractions from a continuous uniform distribution spanning from 0 up to 40 years.*

It can be easily seen that the key nuclide Co-60 has the largest impact on the IRAS. Other important ETMs are Ag-108m and Ti-44. Co-60 is produced via spallation reaction in the element copper but its content can considerably change due to the presence of the element cobalt as an impurity in cathodic copper. The most common reaction to produce Co-60 from natural cobalt is $^{59}Co(n,\gamma)^{60}Co$. Ag-108m is commonly produced by activation of natural silver and Ti-44 by spallation reactions with copper, iron and titanium. Strictly speaking Ti-44 is not an ETM radionuclide because its γ-lines are difficult to detect. Ti-44 however decays into Sc-44 via electron capture (100%) and the daughter element emits a strong γ-photon ($I_\gamma = 99.882\%$) at 1157 keV that can be easily detected via γ-spectrometry. For this reason Ti-44 is included among the ETM radionuclides.

The distinction between DTM and ITM is based on the experience accumulated during the last years of waste characterization at CERN. In particular, we have seen that if a radionuclide contributes for less than 1% to the IRAS its specific activity often lies below the detection capabilities of the instruments used for its quantification. For this reason we decided to fix at 1% the threshold to discriminate between DTM ($C_{IRAS} > 1\%$) and ITM ($C_{IRAS} < 1\%$). We will extensively discuss the quantification of the activity of DTM and ITM radionuclides in Ch. 3.

DTMs (Ni-63 and H-3) and ITM (Fe-55) radionuclides are mainly produced by spallation reaction with copper. The presence of iron as an impurity can increase the production of Fe-55.

The calculations performed using Actiwiz [Theis and Vincke 2012] indicate the production of 45 radionuclides. As shown in Tab. 2.7 only 6 of them have a real impact on the IRAS. The remaining 39 radionuclides have an overall impact below 0.1% on the hazard factor and are not described here. A non-comprehensive list of these radionuclides includes C-14, Na-22, Ar-39, Mn-54, Co-57, Ni-59, Sr-90, Cd-113m, Hg-194, Pb-202 and Bi-207.

□

The previous example showed how we can evaluate the radionuclide inventory for a given material exposed at CERN. Among the radionuclides produced, some can be measured from outside a waste package using non-destructive assay means. These radionuclides are identified as ETMs. Other radionuclides cannot be measured from outside the waste package via non-destructive assay techniques. Of this second category, some radionuclides contribute for more than 1% to the IRAS and are called DTMs. The ones that contribute for less than 1% are called ITMs. Next chapter describes in great detail the instrumental and analytical methods used to quantify the activities of ETM, DTM and ITM radionuclides.

Summary

In this chapter we introduced a methodology to evaluate the radionuclide inventory of VLLW produced at CERN. The methodology is based on extensive Monte Carlo and analytical simulations.

We described the input parameters that must be evaluated when performing these calculations. In particular, we introduced a random vector called scenario **S**. A generic realization of **S** is the unique combination of an elemental composition, a beam energy, a position within the tunnels, an irradiation time and a decay time.

Once a realization is produced it can be used as input for either analytical or Monte Carlo simulations. The simulations so obtained give the list of potential radionuclides produced by activation in that scenario and the so-called correlation factors. The correlation factors allow us to evaluate the relationship between measurable easy-to-measure radionuclides and difficult- or impossible-to-measure radionuclides.

We demonstrated that around 10000 realizations of the scenario should be considered to obtain a robust estimation of the correlation factor. Once the radionuclide inventory is evaluated the produced radionuclides can be classified as either ETM, if they can be measured from outside a waste package via non-destructive assay techniques, or DTMs and ITMs, if they cannot be measured via non-destructive techniques from outside the waste package. We also introduced the concept of contribution to the IRAS to discriminate between DTM and ITM radionuclides.

According to the category they belong, the activity of various radionuclides is quantified using different techniques. The methods used to quantify the activity represent the topic of next chapter.

Chapter 3

Methods to quantify the activity

Each piece, or part, of the whole of nature is always merely an approximation to the complete truth, or the complete truth so far as we know it. In fact, everything we know is only some kind of approximation, because we know that we do not know all the laws as yet. Therefore, things must be learned only to be unlearned again or, more likely, to be corrected.

Richard Feynman
The Feynman Lectures on Physics.

Introduction

Before final disposal, the specific activity of radionuclides in waste packages has to be declared in accordance with limits and criteria derived from safety assessment of the disposal facility. As discussed in Ch. 1 and Ch. 2, some of these nuclides are easy-to-measure (ETM) via non-destructive assay methods, such as γ-spectrometry, from outside the waste package. Other radionuclides are called difficult-to-measure (DTM) because the quantification of their activity requires destructive assay techniques and analytical methods [ISO 2007]. A third category of radionuclides, defined as impossible-to-measure (ITM), is included to account for radionuclides that cannot be measured via non-destructive assay techniques from outside of the waste package and have a very low level of activity, often below the detection capabilities of the instruments used. The activity of this family of radionuclides

is estimated using numerical and analytical methods.

The present chapter is divided into three main sections describing experimental and numerical methods used to quantify the activity of ETM, DTM and ITM radionuclides. The steps of the characterization process described in the present chapter are showed in Fig. 3.1. The complete characterization flux diagram is given in Ch. 1, Fig. 1.10.

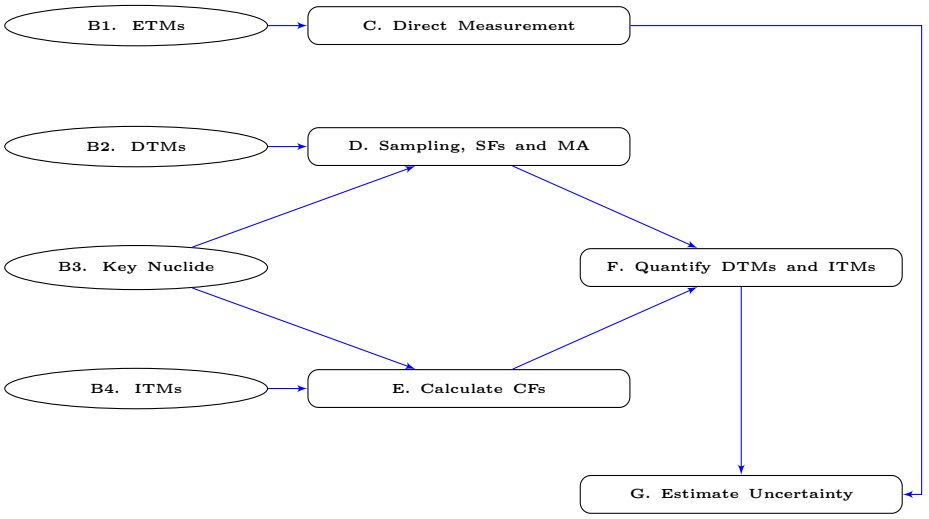


Figure 3.1: Stages of the radiological characterization process to determine the activity of the radionuclides in VLLW. ETMs, DTMs and ITMs stand for easy-, difficult- and impossible-to-measure radionuclides. SF, CF and MA stand for scaling factor, correlation factor and mean activity method.

The first section gives an overview of the instruments used to perform γ-ray spectrometries to quantify the activity of ETM radionuclides (step C of Fig. 3.1). The extended formula to calculate the specific activity of a radionuclide is discussed and the parameters of interest are explained. After describing the requirements of γ-spectrometric systems for waste characterization we discuss the criteria to identify the potential KN for the metallic materials considered in the present thesis (step B3).

The second section begins with the description of the scaling factor (SF) method and the conditions for its applicability. The SF technique is widely applied to the characterization of waste produced in nuclear power plants and is based on the study of the correlation between ETM and DTM radionuclides. The mean activity method (MA) is then introduced to cover the cases in which the scaling factors cannot be applied. The section continues

with a description of basic sampling methods that can be used to choose samples for DTM quantification. The last part of the section is dedicated to the description of the radiochemical methods which are employed to measure the activity of the DTM radionuclides.

The third section discusses the evaluation of the activity of ITM radionuclides via the so-called correlation method (steps E and F of Fig. 3.1). The method lies on intensive numerical calculations, performed using the analytical code Actiwiz [Theis and Vincke 2012], that are used to predict the behaviour of the correlation between the activities of ETM and ITM radionuclides. We use statistical learning to study the relationship between radionuclides' activity and to estimate the importance of the input parameters when evaluating correlation factors.

We end the chapter with a list of DTM and ITM radionuclides which are relevant in the characterization of VLLW produced at CERN.

3.1 Activity quantification of ETM radionuclides

3.1.1 γ-ray spectrometry

Gamma-ray spectrometry is a nuclear instrumental technique which allows the user to identify radionuclides in a sample and to quantify their activity by analysis of the γ photons emitted. In the radiological characterization process, γ-spectrometry is used to quantify the activity of ETM radionuclides on both final waste containers and samples used to evaluate the scaling factors.

A portable γ-spectrometer produced and commercialized by CANBERRA[1], which is commonly employed at CERN, is showed in Fig. 3.2.

The detector of Fig. 3.2 consists of a germanium crystal, which is the active part of the system for γ-detection, a pre-amplifier for the preliminary treatment of the signal collected on the cristal, an amplifier, a power supply and a cooling system [Knoll 2010]. The system is electrically chilled and can work autonomously on batteries for 6 hours without sector

[1] `http://www.canberra.com`, 25 November 2015.

Figure 3.2: Falcon5k is an electrically chilled and portable γ-ray detector. This system is commonly used at CERN to quantify the activity of ETM radionuclides in radioactive waste.

connection. The relative efficiency[2] of the detector is 30% and its energy resolution[3] is $\sim$1.5 keV at 1.33 MeV.

The γ-ray spectrometers used at CERN are characterized. This means that their spatial and angular response is known for the energetic range of photons from 45 keV up to 7 MeV. The energy of interest is however limited to the interval 45 kev - 3 MeV.

Characterized detectors are compulsory in waste characterization for the determination of the specific activity of the ETM radionuclides. In particular, the characterization is useful to estimate the efficiency calibration function $\epsilon(E)$, representing a measure of the probability that a γ-ray of energy E is fully absorbed in the active volume of the detector [L'Annunziata 2012].

A few commercial software packages are available for spectra analysis and efficiency

[2]Many commercial manufacturers specify the photopeak efficiency relative to that of a standard 3 in.$\times$3 in. (7.62 cm$\times$7.62 cm) cylindrical NaI(Tl) scintillation crystal. A source-detector spacing of 25 cm is assumed in both cases for standardization. The relative efficiency is normally specified for the 1.33 MeV γ-ray photopeak of Co-60. The efficiency ratio can be directly measured by simply determining the photopeak area from both detectors, using a Co-60 source [Knoll 2010].

[3]The energy resolution, or FWHM (Full Width at Half Maximum), is the width in keV of a γ peak of energy E at just half of the maximum ordinate of the peak.

calibration function estimation. At CERN the CANBERRA tools called ISOCS and LabSOCS [CANBERRA 2012] are used for efficiency calibration purposes. The efficiency function is used in the general Eq. 3.1 for the calculation of the specific activity a (in Bq/g) of γ-emitting radionuclides [CANBERRA 2009]:

$$a = \frac{S_{net}}{m \cdot \epsilon(E) \cdot I_\gamma(E) \cdot t \cdot K_c \cdot K_w \cdot K_i} \tag{3.1}$$

where:

- S_{net} is the net area under the peak located at the energy E;
- m is the weight of the sample in grams (g);
- $\epsilon(E)$ is the attenuation corrected efficiency function[4];
- $I_\gamma(E)$ is the emission probability of photons of energy E (in %);
- t is the live counting time in seconds (s);
- K_c is the correction factor for the nuclide decay during counting;
- K_w is the correction factor for the nuclide decay from the time the sample was obtained to the start of the collect;
- K_i is the correction factor for decay during the irradiation.

The uncertainty of the specific activity of ETMs, which is needed for the calculation of the overall uncertainty associated to the IRAS (see Sec. 1.3.1), is discussed in Ch. 4, Sec. 4.2.

[4]The attenuation corrected efficiency is given by

$$\epsilon = \epsilon' \cdot e^{\frac{\mu(E)}{\rho}\rho x}$$

where ϵ' is the non-attenuation corrected detection efficiency at the peak energy E, $\frac{\mu(E)}{\rho}$ is the mass attenuation (in cm^2/g) at the energy E and ρx is the average sample mass per unit area [L'Annunziata 2012][CANBERRA 2009].

In Eq. 3.1, S_{net} is the net count area of a peak (assumed to follow a Gaussian distribution with a low-tail) obtained as the difference between a total gross count (S_{gross}) and a background count (S_{bkg}).

When evaluating the specific activity of an object it is common to evaluate its net weight (gross weight minus tare). In the radioactive waste characterization it is often required to communicate to the waste management agency the specific activity of a package (waste plus container). It is therefore not uncommon that, for the evaluation of Eq. 3.1, m represents the gross weight of a waste package.

The efficiency ϵ is a crucial parameter for the evaluation of the activity of ETMs. This term represents the capability of the detection system to collect and measure the photons emitted by the source. The evaluation of the efficiency is made via numerical methods commonly implemented in commercial software. The efficiency is affected by the geometry of the sample/detection system, the nature and density of the sample but also by the activity distribution within the sample to be measured. The efficiency changes with the energy of the incoming photons and is commonly estimated using a multi-gamma source for fixed geometries or a software if the geometry of the system changes routinely.

An example of calibration function obtained for the Falcon5k detector is shown in Fig. 3.3 (top).

The item measured to build the calibration showed is a container of 1.3 m^3 containing iron VLLW. The numerical algorithm used to describe the behaviour of the efficiency is:

$$\ln(\epsilon) = \sum_{i=0}^{n} b_i \cdot (\ln(E))^i \tag{3.2}$$

where b_i is the i^{th} coefficient to be estimated by interpolation, E is the energy and n is the number of points used to evaluate the curve. As it is showed in Fig. 3.3, the efficiency function is generally split in two parts, one for the low energies (from 30 keV to 90-100 keV; in red in the figure) and one for the high energies (above 90-100 keV; in blue in the figure) for achieving a better interpolation (minimization of the residuals).

To ensure the quality of γ-spectrometry results, three more calibration factors must be considered, namely the resolution (FWHM), the tail of the peaks and the energy calibration.

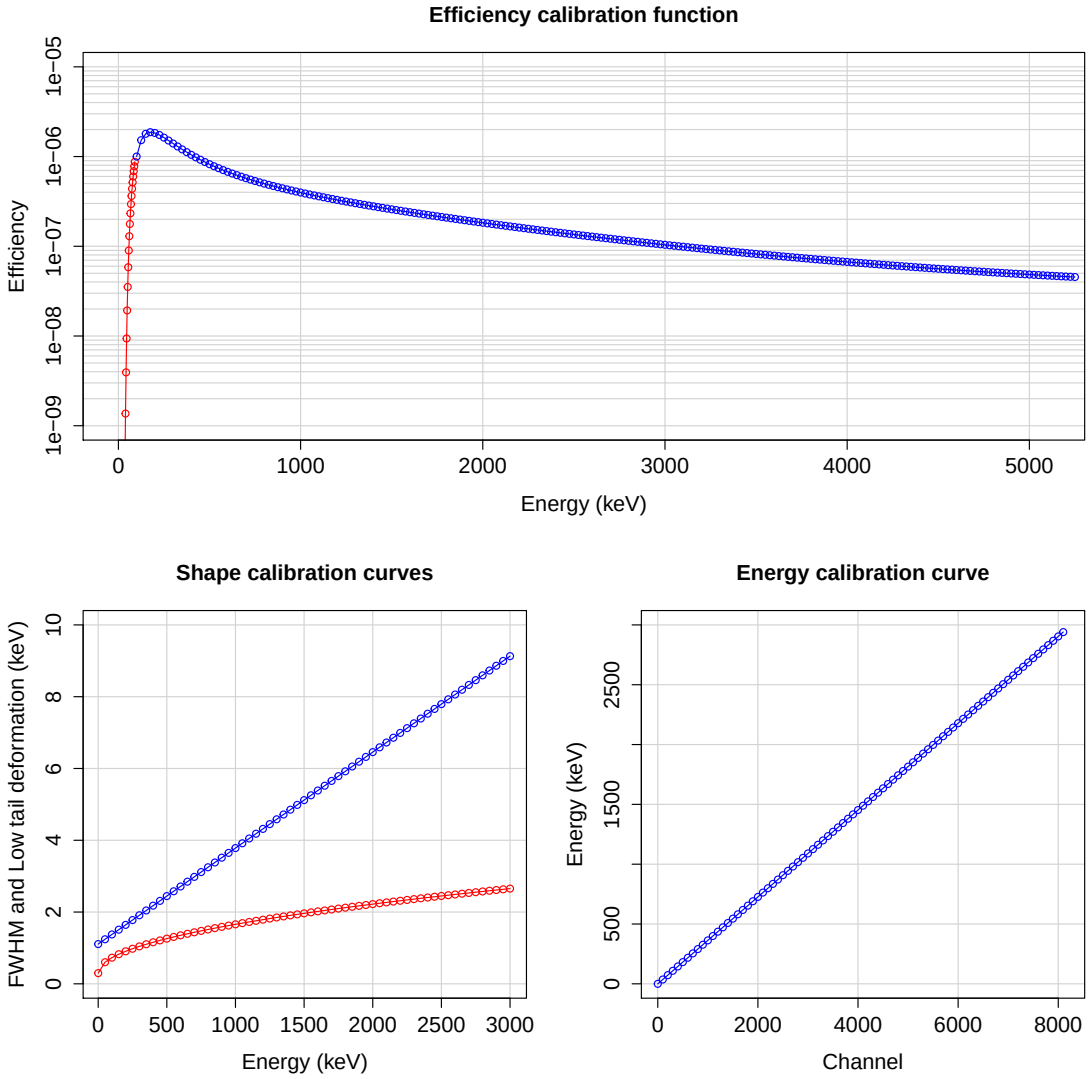


Figure 3.3: Example of calibrations for a standard 1.3 m^3 container filled with pure iron radioactive waste and measured with the Falcon5k. The top plot represents the efficiency calibration function. The bottom-left plot shows the resolution (in red) and the low-tail deformation (in blue). The bottom-right graph represents the energy calibration curve.

As stated before, a photopeak can be modelled as a Gaussian distribution with a low tail. In reality, in a perfect detecting system we should be in presence of a Dirac signal at a fixed energy. Because of the electronic noise, the statistical nature of the decay and the imperfect collection of the charges, a photopeak results in a fairly normal distribution over multiple channels of a multi-channel analyser (MCA). The MCA is the electronic

component of the detection system that sorts the electric charges created during the energy deposition of the photons in the crystal and discriminates them according to their height, finally generating the so-called γ-spectrum. The FWHM increases with the energy of the incoming photon according to a law of the type:

$$FWHM \propto \sqrt{E}. \tag{3.3}$$

The FWHM of the peaks generated by a γ-detector deteriorates with the time but experience demonstrate that a detector can give quality data for periods spanning over 15 to 20 years. An example of resolution for the Falkon5k detector is given in Fig. 3.3 (bottom left, red line).

Another effect that can have an impact on the quality of a spectrum is the so-called "pile-up". This effect is due to an interference between pulses when the counting rates are high. In practice, when such an effect is present, the tails of a Gaussian peak can be affected by a distortion [Knoll 2010]. Moreover, this effect increases with the energy of the incoming photon according to a law of the type:

$$\text{Tail} \propto E. \tag{3.4}$$

An example of low-tail deformation function is showed in Fig. 3.3 (bottom left, blue line).

Finally, to correctly identify the energies of the photons it is necessary to establish a correspondence between the channels of the MCA and the energy of the incoming photon. This energy calibration is checked routinely and performed at the first utilization of the instrument. An energy recalibration can be required if the environmental conditions and the use of the instrument change.

The energy calibration function has the form:

$$E = \sum_{i=0}^{n} c_i \cdot Ch^i \tag{3.5}$$

where c_i are the coefficients to be determined at the moment of calibration with standard

sources and Ch is the channel of the MCA. In general the sum up to $n = 2$ is a good approximation for the true relationship. An example of energy calibration function is represented on the bottom-right graph of Fig. 3.3.

In Eq. 3.1 the γ intensity, or probability emission, represents the percentage of photons emitted per primary decay. The values for the emission probabilities are regularly updated and published by reference institutes or working groups via on-line accessible database.

The counting time t is the effective time during which the detection system is able to detect and process the information associated to an incoming photon. Due to the amount of signals received by a detector, it can happen that the system is not able to process a new signal for a given amount of time (called dead-time). A correction for dead-time must therefore be applied to calculate the precise counting time t used in Eq. 3.1.

Finally, for waste characterization purposes, the correction factor $K = K_c \cdot K_w \cdot K_i$ can be neglected. This is mainly due to the short counting time of ETM's activity in a package (usually 15 minutes) compared to the half-life of the ETMs and to the fact that we deal with historical waste (short-lived nuclides have long disappeared when the measurement starts). Moreover, we are interested about the specific activity of the radionuclides in a waste package as obtained from measurements at a given time and the radiological history of the waste does not affect the measurements.

A complete description of the parameters here presented can be found in references [Knoll 2010][L'Annunziata 2012][CANBERRA 2009][Gilmore 2008].

3.1.2 Practical aspects in gamma-spectrometry

A few aspects are relevant when performing γ-ray spectrometry to characterize VLLW. One can cite:

- the activity distribution inside a package containing radioactive waste;
- the source-detector configuration;
- the counting time;
- the Minimum Detectable Activity (MDA).

If the activity within a waste container is not homogeneously distributed, care must be taken during the activity calculation process. Non-homogeneous activity distributions can be used when calculating the efficiency calibration function. Examples of activity distributions available on the ISOCS-LabSOCS tool [CANBERRA 2012] are point and volumetric sources within an object, Gaussian distribution and surface distribution of activity (for simulation of contaminated area). The point source can be used to estimate the activity within a package in the worst case scenario of a hot-spot, considering the highest self-absorption within the waste itself.

The source-detector configuration affects the efficiency function because of the solid angle[5] between the sample under measurement and the detector. In particular, the efficiency decreases with the distance source-detector.

A very common waste package used at CERN for VLLW conditioning is a steel box of 132×102×98 cm^3. Fig. 3.4 shows a batch of standard containers ready to be measured via γ-ray spectrometry.

When this type of container is used, a good compromise in terms of distance source-detector is 50-to-75 cm. This range of distances allows us to perform measurements with a satisfactory efficiency and, consequently, a counting time short enough to reach quickly the desired accuracy (the counting time is often set to be 15 minutes). A common hypothesis in setting the measurement conditions is that the chosen counting time must be long enough so that the MDAs of each ETM are below the Declaration Thresholds (DTs; see Sec. 1.3.1). Practical experience indicates that a counting time of 15 minutes ensures detection levels 1-to-2 orders of magnitude below the DT.

To decrease the heterogeneity associated with the activity distribution in a waste package, an operational dose-rate criterion is set when choosing the items that will be

[5] In γ-ray spectrometry the solid angle Ω is defined by an integral, over the detector surface that faces the source, of the form:

$$\Omega = \int_A \frac{\cos\alpha}{r^2} dA$$

where r represents the distance between the source and a surface element dA, and α is the angle between the normal to the surface element and the source direction. If the volume of the source is not negligible, then a second integral must be carried out over all volume elements of the source [Knoll 2010].

Figure 3.4: A batch of 1.3 m^3 packages containing metallic VLLW. Credits: B. Cellerier, CERN.

conditioned within that package. In particular, as far as possible, items with similar dose-rate and radiological history are conditioned together. Moreover, multiple measurements are performed on different sides of the same waste package and an average activity is finally calculated as best estimator of the true mean activity. We have demonstrated that 2 measurements performed on the bigger face of a waste container are enough to completely characterize a waste package [Magistris and Zaffora 2015]. We estimated a loss of accuracy with respect to a 4-sides characterization below 1%.

A last parameter of practical interest is the Minimum Detectable Activity (MDA). The calculation of the MDA (in Bq/g) for a given nuclide, at the 95% confidence level, is based on Currie's derivation [Currie 1968], with one formulation being [Gilmore 2008] [De Geer 2004] [Kirkpatrick et al. 2013]:

$$MDA = \frac{f(\sigma_{bkg})}{t \cdot \epsilon(E) \cdot I_\gamma \cdot m \cdot K} \tag{3.6}$$

where $f(\sigma_{bkg})$ is a function of the standard deviation of the background collected during the time t (in seconds) over the energy range of interest. The other symbols are explained with Eq. 3.1.

This formulation yields the smallest level of activity which can be detected with 95% confidence, while also having 95% confidence that radioactivity is not detected falsely from

a null sample.

From Eq. 3.6, we can see that the MDA depends on many factors. As a general rule, the MDA decreases with increasing counting time, which means that a longer measurement offers the opportunity to observe lower activities. The MDA depends directly on the level of the background (which is essentially the natural radioactivity present in the measurement site together with the artificial one, when the sample is not present). It is then preferable to perform measurements in low level backgrounds. The consequences of a low background are lower MDAs and spectra which are simpler to analyse. A complete mathematical treatment of this topic can be found in references [Knoll 2010][L'Annunziata 2012][CANBERRA 2009][Gilmore 2008][Currie 1968] and [De Geer 2004].

A new standard for the determination and interpretation of characteristic limits for radioactivity measurements was published in 2010 (ISO 11929 [ISO 2010]). IAEA released in 2017 a technical report in which the new standard is interpreted and discussed [IAEA 2017]. The new standard is not presently applied at CERN when performing calculations related to γ-spectrometry and, for this reason, in the sections and chapters that follow we will refer to the MDA when discussing about detection capabilities of nuclear instruments and techniques.

3.1.3 Selection of potential key nuclides for metallic waste

The characteristics necessary for an ETM to be identified as a key nuclide are introduced in Sec. 1.4 and are discussed in references [ISO 2007] and [IAEA 2009b]. We briefly recall here that a key nuclide must have a relatively long half-life (compared to the characterization process), its activity must be correlated to that of DTM and ITM radionuclides and it must be easily measurable using non-destructive techniques from outside a waste package.

Except for a very limited number of cases in which a correlation based on production mechanisms can be established, for historical waste the scaling factor method relies on a purely empirical relationship. The specific activity of the difficult-to-measure radionuclide can be inferred reliably from the activity of the key nuclide as long as a correlation, based on experimental data, can be confirmed for the pair DTM/KN.

The waste target population of the present thesis is legacy metallic waste. A special focus is given to steel, aluminium and copper because these materials represent the vast majority of material types used to build accelerator components. When activating metals at CERN a long list of radionuclides is produced. This is mainly do to the high number of trace elements present on materials but also to the high number of different grades used.

We performed an extensive series of activation studies to identify the main radionuclides and their contribution to the IRAS when metals are irradiated at CERN. We considered more than 40 material grades for over 2.35 million activation scenarios which are considered as representative for CERN's accelerator complex. Tab. 3.1 presents the list of potential key nuclides obtained from the simulation. The key nuclides presently used for characterization are given in bold.

Table 3.1: List of potential key nuclides for metallic VLLW characterization. In bold are given the key nuclides presently used for these material families.

Material	**KNs**	**$T_{1/2}$ (y)**	**Main γ-emitters**
Steel	Na-22	2.603	1275 keV
	Ti-44	58.9	1157 keV (from Sc-44)
	Co-60	5.2711	1173 keV, 1332 keV
Copper	Na-22	2.603	1275 keV
	Ti-44	58.9	1157 keV (from Sc-44)
	Co-60	5.2711	1173 keV, 1332 keV
	Rh-101	3.3	127 keV, 198 keV, 325 keV
	Sb-125	2.7586	428 keV, 601 keV, 636 keV
Aluminium	**Na-22**	2.603	1275 keV
	Ti-44	58.9	1157 keV (from Sc-44)
	Co-60	5.2711	1173 keV, 1332 keV

The radionuclides Na-22, Ti-44 and Co-60 must be considered as relevant independently from the material under investigation. Ti-44, whose main γ-rays (68 keV and 78 keV) are difficult to use to estimate the activity due to multiple interferences with natural occurring radionuclides, is quantified via measurement of its daughter's γ-ray, the Sc-44 (E_γ=1157 keV).

For very special and limited grades of copper, also Rh-101 and Sb-125 can be considered as key nuclides but a specific investigation of the chemical composition must be performed to quantify traces responsible for their production.

To identify an ETM radionuclide as KN a correlation with DTM radionuclides must

be demonstrated. Only when a correlation can be established the KN can be identified for a batch of waste or chemical composition. If no correlation is found, the scaling factor method cannot be applied to estimate the activity of DTM nuclides and an alternative method of characterization should be used. We discuss these topics in the next section.

3.2 Activity quantification of DTM radionuclides

3.2.1 The scaling factor method

The scaling factor (SF) technique for radioactive waste characterization is described in references [ISO 2007] and [IAEA 2009b]. The international standard ISO 21238 [ISO 2007] outlines the theoretical justification of the SF method, describes the applicability conditions and presents the mathematical methods for the evaluation of scaling factors. The IAEA technical report given in [IAEA 2009b] details how different countries apply and calculate SFs when characterizing radioactive waste produced in nuclear power plants.

In this section, starting from the bibliographic references previously given, we describe the method and discuss its applicability to radioactive waste produced at CERN.

Basis for the evaluation of scaling factors

The systematic measurement of the activity of DTM radionuclides on a radioactive waste population is often practically not feasible because it requires long and expensive destructive assay techniques involving chemical and radiochemical treatments of the sample. A number of methods were tested and developed at nuclear power plants to quantify DTM's activities [IAEA 2007].

It is possible to identify two major classes of characterization methodologies. The first class relies on the sampling of the waste population and the evaluation of a relationship between easily measurable and difficult-to-measure radionuclides. The second class involves Monte Carlo and analytical calculations that span over a range of potential activation scenarios typical of a given facility. The first family of methods represents the topic of the present section while the second family is later discussed in Sec. 3.3.

The empirical scaling factor method is a technique for evaluating the radioactivity of

DTM nuclides from the radioactivity of a KN, based on the correlations between DTM nuclides and KNs obtained on a selected sample from the population. The activity of KNs and DTM radionuclides is evaluated from experimental measurements that involve both γ-ray spectrometry (for the key nuclide) and radiochemical analysis (for the radionuclides difficult-to-measure).

DTM radionuclides of primary interest are those with very long half-lives or highly radio-toxic. This translates into radionuclides that have a low Declaration Threshold and a low class (see Sec. 1.3.1).

Scaling factors provide a mechanism for estimating the activity of DTM radionuclides in individual waste packages based on limited radiochemical analysis on samples from the bulk waste population. This is achieved by observing the consistent and reproducible relationships between individual nuclides in samples from a batch, which, with reasonable confidence, can be assumed to represent the entire waste population.

Based on the chemical composition of the irradiated material, the accelerator, the location of the material within the tunnel and the radiological history (irradiation and decay times), the specific activity of radionuclides in radioactive waste often shows concentrations extending over several order of magnitude. As an example, Fig. 3.5 shows the theoretical log-normal distributions of Ni-63 and Co-60 specific activities when cathodic copper is irradiated over a large number of random scenarios (10000) at CERN.

Both Ni-63 and Co-60 show a log-normal distribution, which is a derived Gaussian distribution, characterized by a right-skewed tail. This distribution is very common for activity and activity ratios calculated over a large number of activation scenarios.

Ni-63 and Co-60, respectively DTM and KN, have similar production mechanisms when copper is activated at hadron accelerators. Nuclear reactions of the type (n, p) or (γ, pn) are responsible for the production of Ni-63 from natural occurring isotopes of copper, such as Cu-63 and Cu-65. Similar reactions are responsible for the production of Co-60 from copper via the intermediate production of nickel isotopes. Spallation[6] mechanisms can also

[6] A spallation reaction is an interaction in which a light projectile of high kinetic energy interacts with a nucleus and causes the emission of a large number of hadrons (mostly neutrons) or fragments, leaving behind a reaction product which is generally radioactive [Barbier 1969].

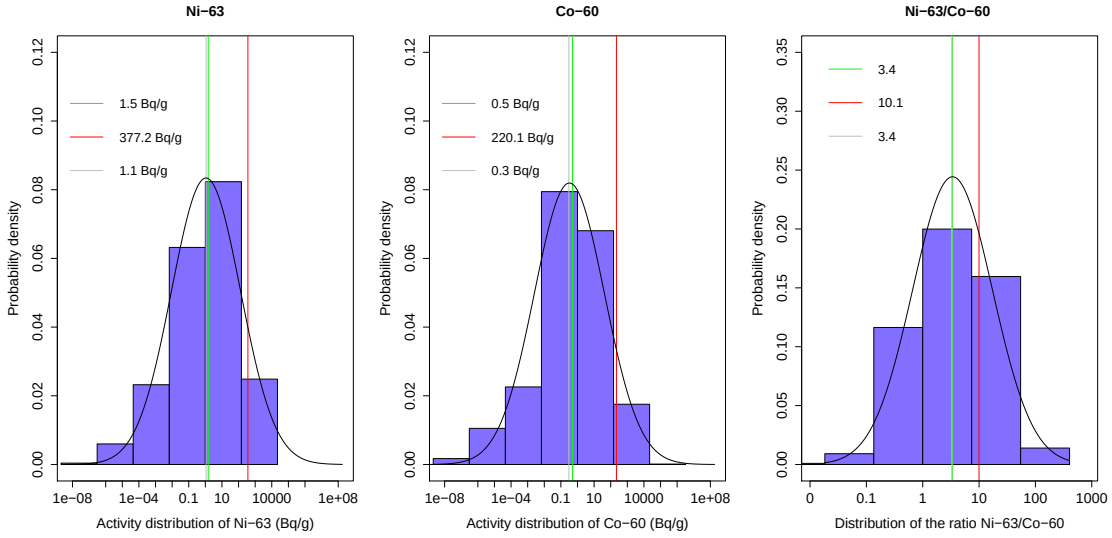


Figure 3.5: Log-normal distributions of Ni-63 and Co-60 specific activities and log-normal distribution of their ratio obtained by simulating the irradiation of cathodic copper at 10000 random CERN scenarios. The average content statistics geometric mean (gray), mean (red) and median (green) are represented together with the normal curve. The x-axis is in log-scale.

be involved.

Fig. 3.5 suggests that also the ratio Ni-63/Co-60 follows a log-normal distribution. Similarity in production mechanism and/or log-normal distribution of the ratio DTM/KN are commonly working hypothesis for the applicability of the scaling factor method. It is however often impossible to identify the mechanisms of production of radionuclides in legacy waste. For these cases, statistical tools can be used to evaluate distribution of activities from data.

SF calculation and applicability

As stated in the previous section statistical methods offer a solution when checking for activity correlations in activated legacy waste.

When a linear relationship exists between the activity of the considered radionuclides, the DTM's activity on a waste package can be predicted by multiplying an appropriately calculated scaling factor and the measured specific activity of the KN in that package [ISO

2007].

Two common methods adopted at nuclear power plants to evaluate SFs are linear models and the geometric mean SF [ISO 2007] [IAEA 2009b]. Independently from the method used, the predicted specific activity $\hat{a}_{DTM_{i,j}}$ of the i^{th} DTM in the waste package j is obtained as follows [Zaffora et al. 2016]:

$$\hat{a}_{DTM_{i,j}} = SF \times a_{KN_j} \tag{3.7}$$

where a_{KN_j} is the measured specific activity of the KN in the package j. For completeness, the KN should have a second subscript k to account for multiple KNs. Such is the case, for instance, when a mix of materials is present within the package or when we can establish correlations between DTMs and multiple KNs.

The generic equation of the linear model can be written as follows:

$$\hat{a}_{DTM} = \hat{\beta}_0 + \hat{\beta}_1 \times a_{KN} \tag{3.8}$$

where $\hat{\beta}_0$ and $\hat{\beta}_1$ are the intercept and the slope of the model. In Eq. 3.8 the intercept is often set to zero. This is a common hypothesis in waste characterization indicating that either both KN and DTM have an activity different from zero or both have an activity equal to zero. Of course this hypothesis may not be true especially when the radionuclides have a large difference in half-life and one of them is completely decayed.

The slope $\hat{\beta}_1$ corresponds to the evaluated scaling factor from the linear model built using n values of activity from a collected sample. We can use common statistical tools, such as the coefficient of determination R^2, hypothesis testing (p-value) and F-statistics, to evaluate the robustness of the linear model so built. Limits and operational values for such quantities can be found in references [IAEA 2007] [IAEA 2009b].

When we collect a sample of n units to evaluate scaling factors, the SF of the i^{th} unit is given by the ratio of the activities of the DTM and the KN:

$$SF_i = \frac{a_{DTM_i}}{a_{KN_i}}. \tag{3.9}$$

When discussing of correlation factors in Ch. 2, we have seen that activity ratios often show log-normal distributions. The geometric mean is a good central estimator for such a family of distributions. As for the correlation factors (see Ch. 2, Eq. 2.10) we can evaluate the specific activity of DTM radionuclides using the geometric scaling factor $\overline{G}_{SF}$ which is defined as follows:

$$\overline{G}_{SF} = exp\left(\frac{1}{n}\sum_{i=1}^{n}\ln(SF_i)\right) = \left(\prod_{i=1}^{n} SF_i\right)^{\frac{1}{n}}. \tag{3.10}$$

where SF_i is given by Eq. 3.9. The only difference between the calculation of scaling and correlation factors lies on that SFs are evaluated from experimental data while CFs are obtained from theoretical activation studies. In practice, correlation factors are preliminary estimation of the scaling factors or, to be more precise, of the potential distribution of the scaling factors.

The linearity between activities of DTMs and KNs, even if often seen, cannot be systematically demonstrated. In such a case, the reference standards suggest the use of alternative methods involving non-linear regression or interpolation. The interested reader can find the detail of these methods in reference [ISO 2007].

Reference [IAEA 2009b] suggests to test the applicability of the scaling factor method via a number of statistical criteria which are based on the experience accumulated by nuclear power plants. A first test consists of demonstrating that 95% of the experimental SF_i values are within the interval:

$$\frac{\overline{G}_{SF}}{10} \leq (SF)_i \leq \overline{\mathrm{G}}_{\mathrm{SF}} \cdot 10. \tag{3.11}$$

A second applicability criterion states that the dispersion $D_{2\sigma}$ around the geometric mean at a 95% confidence level should be $D_{2\sigma} \leq 10$. As for the correlation factors, the dispersion of scaling factors can be calculated as follows:

$$\mathrm{D} = exp\left(\sqrt{\frac{\sum_{i=1}^{n}[\ln(SF_i) - \ln(\overline{SF})]^2}{n-1}}\right) \tag{3.12}$$

where $\overline{SF}$ is the average scaling factor.

A last applicability test suggested by [IAEA 2009b] indicates that the correlation coefficient r between DTM and KN activities must be above 0.5. We discussed earlier in this section that common hypothesis tests and statistics (namely the p-value and the F-statistic) can be used to study linear models.

In some cases a robust correlation cannot be established among radionuclides. Reference [IAEA 2009b] suggests that the so-called "mean activity method" can be used instead. This is the topic that we discuss in the next session.

3.2.2 Mean activity method and bootstrap

In absence of a reliable technique to establish a correlation between DTMs and KNs, the guide of reference [IAEA 2009b] suggests to apply a different technique to evaluate DTM's activity. This method is based on the collection of n samples from the waste population and on the quantification of the activities performed via destructive techniques. For a given waste population n activities of DTMs are therefore available.

The mean activity method (MA) consists of calculating the arithmetic average of all the specific activities of a given DTM on a batch of waste from samples, including activities which are below the MDA. This mean activity is systematically attributed to each waste package of the waste batch. Alternative approaches to the mean activity method includes a different choice of the average content statistics. We could for instance choose the median, the upper limit of the confidence interval around the mean or the 75^{th} percentile. This choice should however be motivated by the characterization philosophy adopted. For instance, if we would like to limit the impact of extreme values we could choose the median and if we would like to be conservative on our estimation we could choose the upper bound of a confidence level. At CERN the philosophy presently adopted is conservative and the mean activity plus two times the standard error of the mean is used when applying the MA method for the characterization of VLLW.

The use of the arithmetic mean or the upper bound of the confidence interval can however be overly conservative, especially when the distribution of the parameter of interest is strongly skewed or if our sample includes outliers. Such an approach could for instance be used when testing worst case scenarios.

A second consideration rises from the use of MDA values as substitute of specific activities when the activity is below this threshold. A vast literature exists around this topic and guidelines can be found in references [EPA 1994] [Aitchison 1955] [Coghran and Egeghy 2003]. In this context we simply want to outline that it seems reasonable to include values below the MDA to calculate the mean activity because we want to estimate an average content of the entire waste population. Values below the MDA carry therefore an important information from the sampling units they were collected from.

In the present thesis, as a complement of the mean activity method, we introduce a new approach to estimate a mean activity value of DTMs [Zaffora et al. 2016]. This approach lies on an intensive numerical technique known as bootstrap. The bootstrap is a statistical tool that can be used for a variety of applications including uncertainty quantification and distribution parameters evaluation. A complete treatment of this technique can be found in references [Efron and Tibshirani 1993] [James et al. 2017] [Hastie et al. 2009].

The bootstrap emulates the process of obtaining new sample sets, so that the variability of a parameter can be estimated without generating additional samples. Rather than obtaining several independent datasets from the population, we obtain several distinct datasets by repeatedly sampling observations from the original set. This technique corresponds to a sampling with replacement, which means that the same observation can occur more than once.

Bootstrap can be particularly useful when a limited number of samples is collected and the estimation of the variability is difficult. This is for example the case of samples collected via judgemental or biased sampling or when the budget for sampling is very limited and the final number of samples is small.

If a dataset Z containing a limited number of observations is available, n observations from the dataset are randomly selected in order to produce a bootstrap dataset, Z^{*1}. The bootstrap dataset Z^{*1} is used to produce a new bootstrap estimate for the population parameter α, called $\hat{\alpha}^{*1}$. This procedure is repeated B times for some large value of B, in order to produce B different bootstrap data sets, $Z^{*1}, Z^{*2}, ..., Z^{*B}$, and B corresponding α estimates, $\hat{\alpha}^{*1}, \hat{\alpha}^{*2}, ..., \hat{\alpha}^{*B}$. From the data so obtained we can build an estimate distribution of the parameter of interest, calculate a robust central tendency estimator and its variance

and the standard error of these bootstrap estimates. The standard error of the estimates is given by:

$$SE_B(\hat{\alpha}) = \sqrt{\frac{1}{B-1}\sum_{r=1}^{B}\left(\hat{\alpha}^{*r} - \frac{1}{B}\sum_{r'=1}^{B}\hat{\alpha}^{*r'}\right)^2}. \tag{3.13}$$

SE_B is an estimate of the standard error of $\hat{\alpha}$ estimated from the original dataset.

Example 5 *To illustrate an application of the bootstrap technique, a population of 87 drums containing shredded copper cables was considered. 87 representative samples were collected, one from each drum, and the specific activity of Ni-63, a(Ni-63), was measured.*

The bootstrap approach is illustrated in the central panel of Fig. 3.6, which displays a histogram of 1000 bootstrap estimates of the activity of Ni-63, each computed using a distinct bootstrap dataset of 25 observations. The histogram was created using a single dataset, and hence could be created using real data.

Only values above the MDA were considered for this example but MDAs can be included when estimating mean bootstrap activities.

The histogram in the central panel looks very similar to the left-hand panel which displays the idealized histogram of the estimates of the activity of Ni-63 obtained by generating 1000 simulated datasets (n = 25) from the true population. The bootstrap estimate standard deviation of â(Ni-63) is 0.064, very close to the estimate of 0.060 obtained using 1000 simulated datasets.

The right-hand panel displays the information in the central and left panels in a different way, via boxplots of the estimates for a(Ni-63) obtained by generating 1000 simulated datasets from the true population and using the bootstrap approach. The boxplots are quite similar to each other, indicating that the bootstrap approach can be used to effectively estimate the variability associated with a(Ni-63).

□

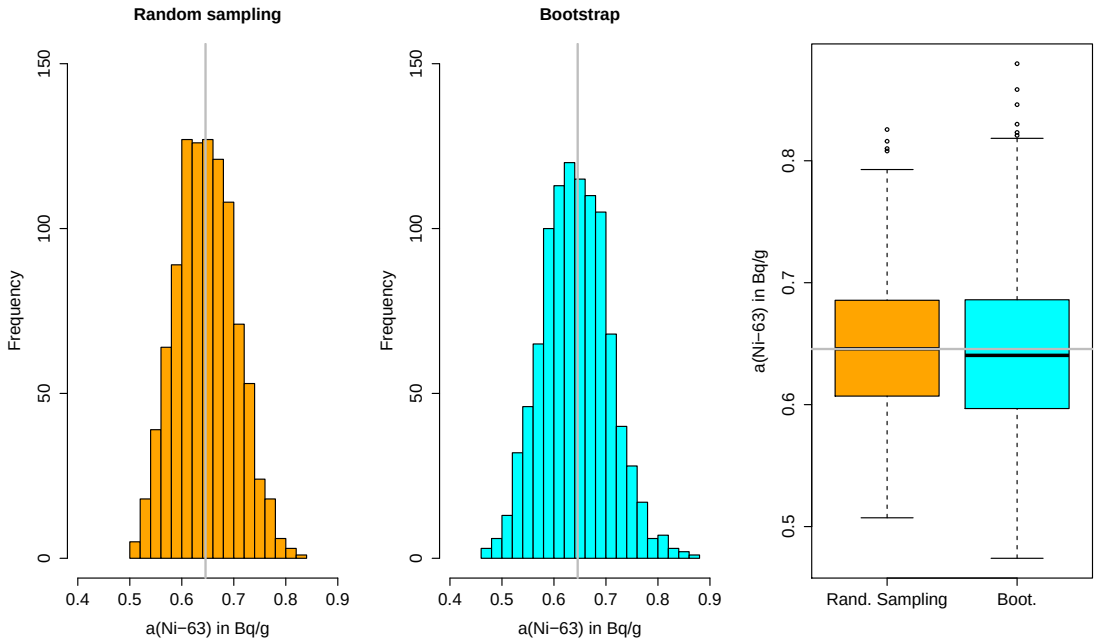


Figure 3.6: Left: a histogram of the estimates of a(Ni-63) obtained by generating 1000 simulated datasets from the true population. Center: a histogram of the estimates of a(Ni-63) obtained from 1000 bootstrap samples from a single dataset. Right: the estimates of a(Ni-63) displayed in the left and center panels are shown as boxplots. In each panel, the gray line indicates the true value of a(Ni-63). This plot is adapted from [James et al. 2017].

3.2.3 Probability and authoritative sampling

The characterization of a radioactive waste population is often based on the analysis of statistical parameters such as the mean, the variance and the standard deviation of the specific activities of radionuclides. The robustness of the statistical parameters used to describe a population depends on the quality of the methods applied to gather the data from which conclusions can be inferred.

In data analysis process, sampling designs provide the methodology for the collection of data and their analysis. In this section we shortly present few categories of sampling designs. As a general classification, sampling designs can be of two types:

- probabilistic, if all parts of the waste under study have a known probability of being included in the sample;

- authoritative, if the knowledge of the analyst is used to infer information over a population in a non-probability context.

The mathematical and statistical treatment of the principal sampling techniques is described in classical books written by Cochran [Cochran 1977], Gilbert [Gilbert 1987] and Myers [Myers 1997]. Useful information can also be found in [EPA 2002] and [Pitard 1993].

In this section, after a short introduction covering the principle steps of a sample plan, we shortly present some common probability and authoritative sampling designs that are of interest for the thesis. The last part of the section introduces the concept of heterogeneity [Pitard 1993] [Gy 1979] [Gy 1998] and some common practices that should be adopted when sampling materials.

Steps of a sampling design

As it is specified by Cochran in [Cochran 1977], it is often impossible to collect "census" data, which is data of all the units in a population. Sampling techniques have, with respect of census studies, many advantages such as, for example, a reduced cost and a greater speed.

For the purpose of the present thesis, the data life cycle for radioactive waste characterization is structured as in references [EPA 1994] [EPA 2000] and comprises three steps called "planning", "implementation" and "assessment".

The planning phase can be divided in the following sub-steps:

- aim of the study: a clear statement of the purpose of the study (for example, evaluation of the specific activity of DTM radionuclide for SF calculation);
- population to be sampled: in waste studies a population is often a batch of waste sharing common characteristics such as the origin, the chemical composition and the radiological history;
- data to be collected: all the data collected must be relevant for the purpose of the survey and must be traced according the quality management system in place;

- degree of accuracy: the degree of accuracy must be defined by the analyst that will use the data and must be thought as the amount of error that can be tolerated in estimates, consistent with making good decisions;
- methods of measurements: in radioactive waste characterization this step refers to the methodologies used to quantify the parameters of interest such as, for example, the activity of α-, β- and γ-emitters, the chemical composition and the dose rate;
- the frame: before selecting the sample, the population must be divided into parts that are called "sampling units" (drums, containers, unitary pieces, etc.); the list of sampling units is called "frame".

After the planning phase is concluded, the implementation starts. We can identify the following stages:

- selection of the sample: there exists a large variety of plans by which samples can be selected including simple random sampling, systematic sampling, stratified sampling, bias and judgemental sampling;
- organization of the field work: in this stage of the survey many organizational challenges arise and must be taken into account by, for example, including on the survey procedures for early checks of the quality of data or the treatment of "non-response", that is the failure to obtain information from certain of the units on the sample (example of values below MDAs).

Finally the assessment phase can starts. In this step the summary of the data is performed and the lessons learned are stated. The two following steps describe the last phase of the data life cycle:

- summary and analysis of the data: this stage allows the analyst to quantify the parameters of interest and their uncertainty;
- information gained for future surveys: any completed sampling campaign is a guide to improve future sampling, both for process knowledge (time, constraint and manpower) and cost involved.

Specific parts of the sampling design will be illustrated in the following sections.

Probability sampling

In probability sampling we are mainly interested in evaluating a principal statistical parameter (such as the mean, the variance and/or the standard deviation) of a waste population, in calculating the number of samples to collect for a given cost or for a pre-specified error and in optimizing the allocation of the samples in the sub-groups within the population. The sub-groups in a waste population often are material families, such as aluminium, copper and steel.

We recall here the results valid for some statistical sampling methods, including simple random, systematic and stratified sampling. For the reader interested in the detailed mathematical treatment of these designs, references are given at the beginning of Sec. 3.2.3.

An application of the various techniques to the radiological characterization of radioactive waste produced at CERN is given in reference [Zaffora et al. 2016].

Simple random sampling

In simple random sampling each member of the population has an equal probability of being included in the sample. Mathematically, simple random sampling is a method of selecting n units out of the N available in the population such that every one of the C_N^n distinct samples has an equal chance of being drawn [Cochran 1977]. In practice, the units of the population are numbered from 1 to N. A series of random numbers between 1 and N is drawn. At any draw, the process must give an equal chance of selection to any number in the population not already drawn. The units that bear these n numbers constitute the sample. Finally, after the collection of the samples, population mean and variance can be estimated using the calculated sample mean and variance.

At CERN simple random sampling for waste characterization was introduced recently (2017). This technique is unbiased only when the entire waste population is accessible. For this reason simple random sampling is performed when all the waste items of a batch of waste must be sorted and dismantled. In this context, a sample per waste package and per material type is collected. The final number of samples to be measured via γ-spectrometry

and radiochemical analyses is established according to the size of the waste batch (on average one sample per ton of waste is measured). At least 20 samples per waste batch are collected.

A sample collected using simple random sampling is representative of the waste population. The scaling factors obtained from such a data collection design are robust estimators of the real unknown scaling factors.

Systematic sampling

Systematic sampling is a statistical process that allows the analyst to choose n samples over a population of N units, with samples spaced by a factor k. If the N units of the population are numbered between 1 and N and n samples must be collected, k is calculated as the ratio N/n. A random sample between 1 and k and every k^{th} unit thereafter are taken.

A second method, suggested by Madow in [Madow 1953], implies taking as starting number not a random number but the value of $(k+1)/2$ if k is odd and either $k/2$ or $(k+2)/2$ if k is even. In any case, the calculations carried out to quantify the estimators that describe the population are effected by the chosen starting point.

At CERN systematic sampling is commonly employed when waste items with similar characteristics must be sampled and they are stored in such a way that they are completely accessible. Examples of waste items commonly sampled using systematic sampling are pipes used for cooling systems, unitary massive waste items made of the same material, originating from identical locations and dismantled at the same time, and accelerator components which are made of the same materials and were installed at the same locations with respect to the beam line, such as vacuum chambers.

Stratified random sampling

In stratified random sampling the population of N units (such as waste containers used for temporary storage) is divided into non overlapping subgroups, called strata.

In the characterization process of radioactive waste, strata are based on specific selection

parameters such as, for example, the material, the γ activity, the dose rate and the radiological history. Samples from strata are obtained applying the rules of either simple random or systematic sampling to each stratum.

A common stratification strategy used at CERN is based on material families. For instance, if a batch of metallic waste is made of aluminium, copper and steel, three strata are identified (one per material type) and the size of the sample collected in a stratum is proportional to the size of the stratum.

Composite sampling

To conclude this summary of statistical sampling designs, we would like to cite composite sampling. Composite sampling consists of mixing together multiple increments coming from the same waste item (or package) into a unique sample. The unique sample so collected will reflect an average characteristic of the waste item under study.

Composite sampling is commonly employed at CERN when collecting samples from an heterogeneous waste population made of the same material type. The case study described in Ch. 5 will show an example of application of composite sampling applied to the characterization of activated shredded copper from cables.

Authoritative sampling

Authoritative sampling is a non-statistical sampling design because it does not assign an equal probability of being sampled to all portions of the population. The validity of the data gathered with authoritative sampling is dependent on the knowledge of the sampler and, although valid data sometimes can be obtained, it is not recommended for the characterization of waste when the parameter of interest (like the average specific activity) is near the activity limit.

Authoritative sampling may be appropriate under circumstances such as the following:

- preliminary information is needed about a waste or site to facilitate planning or to gain familiarity with the waste matrix for analytical purposes;
- only a small portions of the population is accessible and judgement is applied to

assess the usefulness of samples drawn from the small portion;

- extremes values are searched for the calculation of the worst scenario;
- sampling is performed to "prove the positive", which means to find values which do not comply with the limits given by authorities.

Among authoritative sampling, the international standards organization ASTM[7] recognizes two types of techniques called "judgemental" and "biased" [ASTM 2014].

The goal of judgemental sampling is to use process or site knowledge to choose one or more sampling locations to represent the average specific activities or typical properties.

Judgemental sampling designs can be cost-effective if the people choosing the sampling locations have sufficient knowledge of the waste. On the other hand, if the people choosing the sampling locations intentionally distort the sampling by a prejudiced selection, or if their knowledge is wanting, judgemental sampling can lead to incorrect and sometimes very costly decisions.

Accurate and useful data can be generated from judgemental sampling more easily if the population is relatively homogeneous and the existence of any strata and their boundaries is known.

Some disadvantages of judgemental sampling designs are the risk of prejudice when selecting sampling location and the fact that sample variances may be poor estimates of the actual population variance

Biased sampling is used to estimate worst or best case scenarios. The term "biased", as used here, refers to the collection of samples with expected very high or very low concentrations. For example, a sample taken where the dose rate is highest could serve as an estimate of the worst-case γ-activity level within a defined radioactive waste population. It may also be helpful to employ a best-case or both a best-case and worst-case biased sampling approach. For example, if there is a range of parameters characterizing a radioactive waste population and process knowledge can be used to identify the waste likely to have the lowest and highest contamination levels, then these two extremes could

[7] `http://www.astm.org/`, 13 January 2016.

be sampled to help define the extent of the contamination.

Biased sampling, while having the ability to cost-effectively generate information, has similar disadvantages to that of judgemental sampling.

At CERN we have implemented a sampling strategy that combine judgemental, systematic and biased sampling. The reason for making this choice is based on the fact that sometimes, at the beginning of a sampling campaign, dozens tons of waste are not accessible. For these campaigns samples are identified via dose-rate screening. In particular, we look for the highest equivalent dose-rate measurable at the accessible storage locations (this is a combination of biased and judgemental sampling) and we collect a sample per type of waste material (systematic sampling of all the accessible materials) at this location.

The technical report of reference [Magistris and Zaffora 2014] describes authoritative sampling as it is presently applied at CERN.

Sampling, heterogeneity and error minimization

For sampling and analysis techniques to be complete, they must include the use of correct devices and procedures to minimize or control random variability and biases that can be introduced in field sampling, sample transport, sub-sampling, sample preparation and analysis. Sampling error, which is the sum of random variability and bias, can lead to incorrect conclusions, no matter how good the adopted analytical measurements and statistical tools are [Myers 1997].

Sampling theory was developed during the second half of the 20^{th} century by Gy [Gy 1979] [Gy 1998], Pitard [Pitard 1993] and other scientists [Ingamells and Switzer 1973] with the purpose of improving the control of the errors in sampling of materials. They introduced techniques for quantifying the amount of error that can be introduced by the physical sampling process and mathematically described the important concept of heterogeneity.

Gy's theory mainly focuses on minimizing errors during the physical collection of a sample media and should not be confused with the statistical sampling designs [EPA 2002]: Gy's theories facilitate collection of individual samples, while statistical sampling designs

allow the analyst to conduct statistical analyses and make conclusions about the (waste) population.

Sampling theory and heterogeneity's models are complex topics and their detailed description is beyond the purpose of the present thesis. We will however describe few chosen topics because of their importance on the overall process of radioactive waste characterization. In particular, we will present the concepts of random variability, bias and heterogeneity in order to understand common strategies for minimizing sampling errors.

An estimate $\hat{\theta}$ of a parameter θ (such as the mean or the variance) made from measurements of samples always includes some random variability and bias, which can be defined as systematic shift from the true value. Random variability is mainly due the inherent variability of the waste and the bias is due to imprecision in the methods used to collect and analyse the samples [EPA 2002].

Often, errors caused by the sample collection can be greater than preparation, analytical and data handling errors and can dominate the overall uncertainty associated with a characterization study.

Even if the two components of error (random variability and bias) are independent, they are related quantitatively. Errors expressed as variances can be added together to estimate the total variability. Biases can also be added together to estimate overall bias. Theoretically, the sum of all the variances can be added to the sum of all biases and can be expressed as the "mean square error" $MSE(\hat{\theta})$, which is interpreted as the capability of a sample to represent the population.

As indicated in [Myers 1997] and in [EPA 2002], it is often impossible to quantify biases. However, it is strongly recommended to understand the sources and impacts of variability and bias in order to introduce good practices for controlling them in sampling and improve the representativeness of the samples.

Random variability $\sigma^2(\hat{\theta})$ can be defined as follows [EPA 2002]:

$$\sigma^2(\hat{\theta}) = \sigma_p^2(\hat{\theta}) + \sigma_s^2(\hat{\theta}) + \sigma_a^2(\hat{\theta}) \tag{3.14}$$

where $\sigma_p^2(\hat{\theta})$ is the population variability, $\sigma_s^2(\hat{\theta})$ is the sampling and sub-sampling variability

and $\sigma_a^2(\hat{\theta})$ is the analytical variability.

Bias includes sampling bias such as, for example, improper selection and/or use of sampling tools, loss or gain of constituents during sampling, transport, storage and sample preparation but it includes also analytical bias, statistical bias and mistakes.

Finally, we can express the sampling error as follows:

$$MSE(\hat{\theta}) = \sigma^2(\hat{\theta}) + \Big[bias(\hat{\theta}, \theta)\Big]^2 . \tag{3.15}$$

In sampling theory a basic hypothesis is that the material to be sampled is not uniform in its composition or in the distribution of constituents in the mediums. Gy's theory separates heterogeneity into three families called "short-range", "long-range" and "periodic".

Short-range heterogeneity refers to properties of the waste at the sample level. Two different types of short range heterogeneities are identified:

- "constitution heterogeneity" (CH) reflects differences in the composition between individual particles;
- "distribution heterogeneity" (DH) relates to the distribution of the particles in the sampling units and the population.

Constitution heterogeneity is constant and cannot be altered except by particle size reduction (e.g., grinding or crushing the material). Distribution heterogeneity plays an important role in sampling because particles can separate into groups. Distribution heterogeneity can be increased (e.g., by gravitational segregation of particles or liquids) and can be reduced by homogenization (mixing) or by taking many small increments to form a sample (composite sampling). Crushing and mixing methods were applied at CERN in the context of the study performed to characterize activated cables. A detailed discussion will be presented in Ch. 5.

Long-scale heterogeneity reflects local trends and plays an important role in deciding whether to divide the population into smaller internally homogeneous sampling units or to use a stratified sampling design.

Periodic heterogeneity refers to cyclic phenomena found in flowing streams or discharges. This type of heterogeneity is often of no concern for the treatment of legacy radioactive waste. This is mainly due to the size reduction, mixing and compaction performed during the packaging of the waste.

When selecting samples for radioactive waste characterization, batches of waste with similar characteristics are generally identified. At this stage the analysts mainly deals with short range heterogeneities. The description that follows therefore focuses on constitution and distribution heterogeneities.

Let us consider a batch of waste made of discrete units in random order, with a_i being the amount of a generic parameter describing the population (for instance, the specific activity of a radionuclide) in the unit i. The average amount of the parameter a in the population is $\overline{a}$.

The heterogeneity h_i carried by a sample i is defined as:

$$h_i = \frac{(a_i - \overline{a})}{\overline{a}} \cdot \frac{M_i}{\overline{M}} \tag{3.16}$$

where M_i is the mass of the sample i and $\overline{M}$ is the average mass of all the collected samples. The variance of the heterogeneities $s^2(h_i)$ carried by the fragments of a batch of waste corresponds to the constitution heterogeneity of the batch and can be expressed as follows:

$$CH = s^2(h_i) = \frac{1}{n} \sum_i h_i^2 \tag{3.17}$$

where n represents the number of collected samples. In reality, the formula given are true only if each sample represents perfectly the characteristics of the unit from which it is collected. The quantity $\overline{a}$ of Eq. 3.16 can be estimated by the average calculated from the sample.

A similar procedure is used for the definition of the distribution heterogeneity with the difference that for CH the calculation of the heterogeneity is carried out from a sample with respect of the entire population. For DH the heterogeneity is defined for group of fragments which can occupy a large volume. Constitution and distribution heterogeneity can be used to estimate $\sigma_p^2(\hat{\theta})$ of Eq. 3.14.

Gy's sampling theory identifies a list of different errors that can occur in sampling. They are called "fundamental error", "segregation and grouping errors", "delimiting and extracting errors" and "preparation error". Without entering in too much details, it is here presented a list of actions that can be undertaken to reduce these errors [Myers 1997] [EPA 2002] [Pitard 1993].

To reduce the fundamental error, which is caused by difference in the composition of individual particles in the waste, the volume of the sample must be increased and, if possible, the particle size must be reduced and sub-sampling used.

To reduce the grouping and segregation errors, which are due to the existence of clusters in a material, many increments/sub-samples must be taken. The samples should also be homogenized.

To reduce the incremental delimitation/extraction errors the analyst should use sampling devices which are capable of properly delimit and extract all the material that should be included in the sample. For instance, for 3-dimensional waste (drums, containers, massive and unitary pieces) the correct sample is an undisturbed vertical core that captures the full depth of interest.

Finally, to reduce the preparation error, contamination of any type must be prevented, the loss of material must be prevented by a standardized storing procedure, appropriate conditioning during storage must be ensured to avoid chemical/physical transformations and unintentional mistakes must be avoided when labelling, handling and weighing samples through an appropriate system of quality management.

3.2.4 Radiochemical analysis

Radiochemical analyses for CERN's radioactive waste characterization are performed by external laboratories. In this section we shortly describe common radiochemical techniques and measurements methods that are of interest for DTM's activity evaluation.

Radiochemical analyses are performed on samples taken from a waste population to estimate the specific activity of pure β-emitters, low-energy γ-ray and X-ray emitters. These activities are commonly used to estimate the concentration of DTM radionuclides in

the waste packages according to the methods discussed in Sec. 3.2.

The β-emitters are defined as difficult-to-measure radionuclides [ISO 2007] because the quantification of their specific activity requires complex multi-stage techniques involving acid digestion, separation, filtration through resins or columns and measurement. The techniques discussed in this section are mainly related to the measure of H-3, Fe-55 and Ni-63, which are relevant DTM for the radiological characterization of metallic radioactive waste produced at CERN.

A first common step for the quantification of β-emitters in solid samples is an acid digestion. This phase is often necessary to transfer the analytes into a measurable liquid solution and to obtain a complete decomposition of the matrix to be analysed [Patnaik 2004].

Wet chemical digestion utilizing mineral acids (e.g., HCL, HNO_3, HF and H_2SO_4) is carried out in either an open system or in closed vessels. To improve the process the sample can also be heated.

When measuring Fe-55 from pure iron or stainless steel, iron is separated from the prepared sample solution by solvent extraction using, for example, diisopropyl ether. The purified iron fraction is decolourized, and the Fe-55 content is measured by liquid scintillation counting [L'Annunziata 2012]. Fe-55 of known activity is used to determine the counting efficiency. Stable iron measurements, before and after the analytical procedure, is generally used to monitor the chemical recovery.

A similar process is used for the measure of Ni-63 on copper matrix. The stage of separation for Ni-63 in copper differs from the separation of Fe-55 in iron because anion exchange and nickel specific resins are often used for Ni-63 measurements.

Pyrolysis is the method used to measure the total tritium content, which is tritiated water (HTO) and organically-bound tritium (OBT) combined. With this technique, a known mass of sample material is combusted in a two-stage catalytic pyrolyser. The pyrolyser consists of several independently controlled furnaces, through which a silica tube is inserted. The latter half of the silica tube carries an alumina/platinum catalyst, which is maintained in the furnace at 850°C. The heated catalyst oxidises all forms of tritium

to HTO with high efficiency. As the sample temperature is increased, complete oxidation of carbon and organic material is achieved. The outlet is passed into a system of water bubblers where the HTO vapour condenses and the tritium exchanges with the water in the bubblers. An aliquot of the bubbler solution is made alkaline and sodium thiosulphate is added. The water is then distilled at low temperature to remove both radioactive and non-radioactive interferences. An aliquot of the distillate is then measured in a liquid scintillation counter to determine the tritium content. Tritiated water of known activity concentration is used to determine the counting efficiency [L'Annunziata 2012].

The techniques previously described are the basis for the measurements of H-3, Fe-55 and Ni-63 performed by external laboratories under CERN mandate. The measurements of β-emitters used for the calculation of scaling factors in the present thesis were performed by the UK company AMEC[8].

Liquid scintillation counting (LSC) is a common technique used for the quantification of the activity of DTM pure β-emitters. It is based on the property of certain organic materials to emit fluorescent light when interacting with nuclear radiation. The photons emitted by the organic compounds can be converted into an electronic current using a photomultiplier tube. The discussion that follows is based on the classical work of L'Annunziata [L'Annunziata 2012].

The LSC technique involves placing the sample into a scintillation vial and adding a scintillation cocktail. The mixture is homogenized previous to counting. When the energy released by the radioactive decay is absorbed by the organic molecules, light flashes are produced. The intensity of the flashes is proportional to the original nuclear energy dissipated in the fluor cocktail and the number of flashes is proportional to the number of nuclear decays or sample radioactivity.

For most β particles with energy above 100 keV, the counting efficiency is 90-100% but for lower energy emitters (such as H-3) the efficiency is normally in the range 10-60% depending on the degree of quench in the sample.

The quench is defined as the absorption of photons by the sample solution or support.

[8] `http://www.niras.co.uk/`, 27 November 2015.

This phenomenon is not wanted because light flashes produced by fluorescence are lost influencing the correlation between photons emitted and activity in the sample.

It can be possible to identify three different types of quench in LSC, namely physical, chemical and color quenching.

Physical quenching occurs when a barrier forbid contact between radioactive particles and the scintillator solution or when the photons of light are absorbed by solid particles within the vial.

Chemical quenching occurs in presence of any compound that does not have an aromatic structure. Aromatic solvents are used as LSC solvent because of their high density of electrons. When interacting with β particles, aromatic solvents are able to produce a large amount of fluorescence photons.

Finally, color quenching happens when light absorbing compounds are present within the vial and lessen the number of light photons leaving the mixture liquid-solvent.

Various quenching correction methods can be used to correct the quenching effect. The oldest technique for quenching correction is the Internal Standard Method. This technique involves a series of steps for each sample. The first step is to count each sample and obtain an accurate count-per-minute rate (CPM) value for each. Then the samples are removed from the LSC and a known activity of a radionuclide is added to each sample. After the addition of the standard, the samples are recounted to obtain the CPM of the sample plus the internal standard. Once the CPM of the sample and the CPM of the sample plus the internal standard are obtained, the following equation is applied to determine the counting efficiency of the sample:

$$\epsilon = \frac{C_{s+i} - C_s}{D_i} \tag{3.18}$$

where C_{s+i} is the count rate of the sample after addition of the internal standard, C_s is the count rate of the sample without addition of the internal standard and D_i is the disintegration rate of the added aliquot of internal standard. The disintegration rate of the

sample D_s may then be calculated as follows:

$$D_s = \frac{C_s}{\epsilon}. \tag{3.19}$$

More powerful methods for quenching correction, such as the External Standard Method or the Sample Spectrum Characterization Method, are often used in modern LSC systems. Their description is behind the scope of the present thesis and their complete treatment can be found in references [L'Annunziata 2012] and [Ross et al. 1991].

As for the γ-ray spectrometry, also for LSC measurements the MDA is a crucial parameter. In particular, measurements performed using this technique must ensure that the measurable activity from samples is below the VLLW Declaration Thresholds. As described in [L'Annunziata 2012] and [Ross et al. 1991], the MDA can be calculated by the following equation:

$$MDA \simeq \frac{\sqrt{B}}{\epsilon \cdot V \cdot t \cdot K} \tag{3.20}$$

where the MDA is given in Bq/L, B is the total number of counts of the background, V is the volume of the sample in L, t is the counting time in seconds and K is any other factor that may be relevant, such as decay correction, chemical yeald or conversion factor. This equation is valid when sample and background counting times are equal.

For the radionuclides H-3, Fe-55 and Ni-63, which are of interest for waste characterization, common values of the MDA are in the range 0.1-0.5 Bq/L.

3.3 Activity quantification of ITM radionuclides

In Ch. 2 we described the method that can be used to evaluate the radionuclide inventory for a generic material irradiated at CERN. We pointed out that those calculations can also be used to evaluate the specific activity of radionuclides when a direct experimental method cannot be applied.

We recall here that the correlation method is a theoretical approach in which the composition ratios of generic radionuclides and KNs are determined based on the results of

activation calculations which cover the whole range of the elemental composition of the material, the particle spectra and the irradiation conditions [ISO 2013]. In the context of the present thesis, the correlation method is often used for preliminary studies before longer and more expensive analyses which are typical of the SF technique. The correlation method is particularly useful to check for preliminary correlations between DTM radionuclides and KN and to evaluate theoretical correlation factors CF to be used for the evaluation of the activity of ITM radionuclides within waste packages.

The correlation method is fully described in reference [ISO 2013] with a focus on nuclear power plants. We described the methods in Ch. 2 where we discussed how the technique is adapted in order to satisfy CERN's needs.

In practice, a CF is similar to a scaling factor but, instead of being calculated from measurements on samples, it is estimated via analytical calculations or Monte Carlo simulations.

Once a correlation factor is evaluated for a pair ITM/KN, the specific activity $\hat{a}_{ITM_j}$ of an ITM nuclide in package j is estimated as follows:

$$\hat{a}_{ITM_j} = CF \times a_{KN_j} \tag{3.21}$$

where a_{KN_j} is the measured specific activity (in Bq/g) of the key nuclide in the package j and CF is the calculated correlation factor.

When a large number of activation scenarios are taken into account, we must check if for each single scenario a given radionuclide can be above its Declaration Threshold (DT). To do so we can use the proportionality between the activity of each radionuclide and the IRAS. In particular, the ratio of the activity of the i^{th} radionuclide and the IRAS is constant. We can therefore calculate the normalized specific activity a_i^{10} of the radionuclide i when IRAS equals 10 (we recall here that 10 is the maximum allowed IRAS for a waste package to be accepted at the disposal facility as a VLLW) and for a given irradiation scenario as follows:

$$a_i^{10} = \frac{a_i^{IRAS}}{IRAS} \times 10 \tag{3.22}$$

where a_i^{IRAS} is the specific activity of the radionuclide i (in Bq/g/primary/s) for the generic IRAS (for VLLW waste the variable IRAS spans between 0 and 10). Values of a_i^{IRAS} and IRAS are obtained from the analytical calculations performed with Actiwiz or Monte Carlo simulations as discussed in Ch.2.

We finally compare a_i^{10} with the Declaration Threshold of the radionuclide i and include the radionuclide in the inventory if:

$$a_i^{10} \geq DT_i. \tag{3.23}$$

For practical reasons we discriminate between DTM and ITM radionuclides depending on their contribution to the IRAS. In Ch. 2, Sec. 2.1.3 we defined the contribution C_i^{IRAS} of a radionuclide i to the IRAS. If a radionuclide contributes to more than 1% to the IRAS we measure it (direct ETM measurement) or we estimate its activity through sampling and destructive radiochemical methods (DTMs). If a radionuclide contributes less than 1% it is estimated using the correlation method (ITMs). In fact, if the calculated contribution of a radionuclide is below 1%, measurements on sample give back activities which are often below the MDA. For this reason measurements on low-contributing radionuclides is limited to quality measurements and internal consistency validation. A systematic measurement via destructive assay methods would only increase the cost of the characterization strategy without a proportional increase in the quality of the process itself. A simple example of calculation follows.

Example 6 *We simulated the irradiation of a material made of pure iron at the LHC. The material was located in a position close the tunnel wall and the primary beam of protons impacts on magnets. The material was irradiated for 1 years and was left to decay for 10 years. The list of the principal contributors to the IRAS ($C_i^{IRAS} > 0.1\%$) is given in Tab. 3.2.*

The results showed in Tab. 3.2 can be used for a number of purposes. First of all they give us the list of potential radionuclides that have an impact on the IRAS and must be carefully estimated. Other radionuclides are also produced but they are not shown in this example because their impact is negligible.

Table 3.2: Case study of pure iron irradiated for 1 year at the LHC. The material, which was located close to the tunnel wall, is left to decay for 10 years. The contribution to the IRAS is calculated together with the activities for IRAS = 10. The specific activity a_i is obtained from analytical calculations performed with Actiwiz [Theis and Vincke 2012] and is given in Bq/g/primary/second.

Radionuclide	$\mathbf{a_i}$	$\mathbf{a_i/AL_i}$	$\mathbf{C_i^{IRAS}}$	$\mathbf{a_i^{10}}$
Fe-55	$1.39 \cdot 10^{-6}$	$1.39 \cdot 10^{-9}$	94.92%	9492
Mn-54	$5.39 \cdot 10^{-10}$	$5.39 \cdot 10^{-11}$	3.68%	4
H-3	$9.86 \cdot 10^{-9}$	$9.86 \cdot 10^{-12}$	0.67%	67
Ti-44	$9.76 \cdot 10^{-11}$	$9.76 \cdot 10^{-12}$	0.66%	1
Na-22	$8.20 \cdot 10^{-12}$	$8.20 \cdot 10^{-13}$	0.06%	0.1
V-49	$1.39 \cdot 10^{-10}$	$1.39 \cdot 10^{-13}$	0.01%	1
		$IRAS = \sum a_i/AL_i$	$\sum C_i^{IRAS}$	
		$1.47 \cdot 10^{-9}$	100%	

Secondly, comparing the activity that radionuclides should have to be above the Declaration Threshold, if the IRAS was 10, we can predict which radionuclides should potentially be taken into account for the determination of the IRAS.

Finally, we can distinguish between DTM and ITM radionuclides according to the amount they contribute to the IRAS. In this particular case Fe-55 would be classified as a DTM, samples would be taken from the waste population and its activity would be calculated using either a scaling factor or the mean activity method. H-3 and V-49 (decaying respectively via β-minus emission and electronic capture) would be classified as ITMs because their contribution to the IRAS is below 1% and their activity would be estimated using theoretical correlation factors.

The remaining radionuclides (Mn-54, Ti-44 and Na-22) are ETMs and their activity can be directly measured on each package via γ-ray spectrometry.

□

In the following discussion we will present two alternative methods to study and evaluate correlation factors. These techniques make use of the large number of results obtained from simulations and numerical experiments in a statistical learning based approach. We will particularly discuss multiple linear regression and regression trees. The discussion includes parameters importance evaluation, and various techniques to reduce variance, such

as bagging and boosting.

The term statistical learning originates from the fact that a dataset is subdivided into two or more sub-sets. A model is fitted or trained using the data from one of the subsets (called "training set") and its robustness is tested on one or more different sub-sets (called "test datasets"). The guidelines used to implement these statistical learning methods can be found in references [Efron and Tibshirani 1993] and [James et al. 2017].

3.3.1 Multiple linear regression

Multiple linear regression is a statistical technique that can be used to relates a response Y (also called dependent variable) to some predictors X (also called features or independent variables) with the aim of predicting the response for future observations (prediction) or better understanding the relationship between response and the predictors (inference). Multiple linear regression and regression trees (discussed later) are classified under the name of "supervised learning". In supervised learning, for each observation of the predictor measurement x_i $(i = 1, ..., n)$ there is an associated response measurement y_i.

The functional relationship between Y and X can simply be written as:

$$Y = f(X) + \epsilon \tag{3.24}$$

where f is some fixed but unknown function of X and ϵ is a random error term which is independent of X and has mean zero.

Multiple linear regression is a parametric method because we make assumption about the functional form of f. In particular f is linear and has the form:

$$f(X) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_p X_p = \beta_0 + \sum_{i=1}^{n} \beta_i X_i. \tag{3.25}$$

One of the most common procedure to train or fit an appropriate model from the data is called "ordinary least square". When using such modelling techniques care must be paid to overfitting which can be described as taking into account the noise in model construction. Overfitting increases the error associated to prediction and should be avoided.

In the regression settings, the quality of fit can be measured by the mean squared error (MSE):

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{f}(x_i))^2 \tag{3.26}$$

where $\hat{f}(x_i)$ is the prediction that $\hat{f}$ gives for the ith observation.

Multiple linear regression is used when a linear model is to be tested and the number of predictors is higher than one. Such is the case of the correlation factors, whose behaviour depends on the elemental composition, the energy of the accelerator, the position within the machine and the irradiation and decay times (see Ch. 2).

The mathematical treatment of multiple linear regression can be found in a number of classical textbooks such as [James et al. 2017] [Faraway 2004] and [Bishop 2006]. We will describe how such a method can be applied to the evaluation of correlation factors and to parameters importance quantification using a practical example.

The model will include two categorical predictors (energy E and location P) and n numeric predictors (with n-2 chemical elements and irradiation t_i and decay t_c times).

We can write the general expression of the model as follows:

$$CF = \beta_0 + \beta_1 E + \beta_2 P + \beta_3 t_i + \beta_4 t_c + \beta_5 El_1 + ... + \beta_{n-2} El_{n-2}. \tag{3.27}$$

E and P being categorical variables with multiple levels, the coefficient β_1 and β_2 are vectors of dimension 5 and 7 respectively (5 machine energies and 7 possible locations within the tunnels). El_i represents the i^{th} chemical element of the elemental composition. The analysis that follows is performed using the statistical language R [R CORE TEAM 2014].

Example 7 *For this example we consider the activation of cathodic copper. In particular, 10000 scenarios were randomly selected according to the extraction from the random vector* **S** *as discussed in Ch. 2.*

If we want to make predictions about future values of the correlation factor, the dataset should be split in training and test dataset. With this example however we want mainly

understand which features have an impact on the CFs so we will use the entire dataset as is.

A common first step in regression consists of plotting a matrix of the variables versus each other to see if there are visible relationships. This visualization can suggest for instance a transformation on a variable so that its behaviour can be "linearised". As previously discussed, correlation factors often show log-normal underlying shapes and log-transformations are quite useful when building multilinear models.

An example of scatter-plot matrix for a selected number of predictors is showed in Fig. 3.7. In this example we consider the correlation factor of the pair Ni-63/Co-60 in cathodic copper.

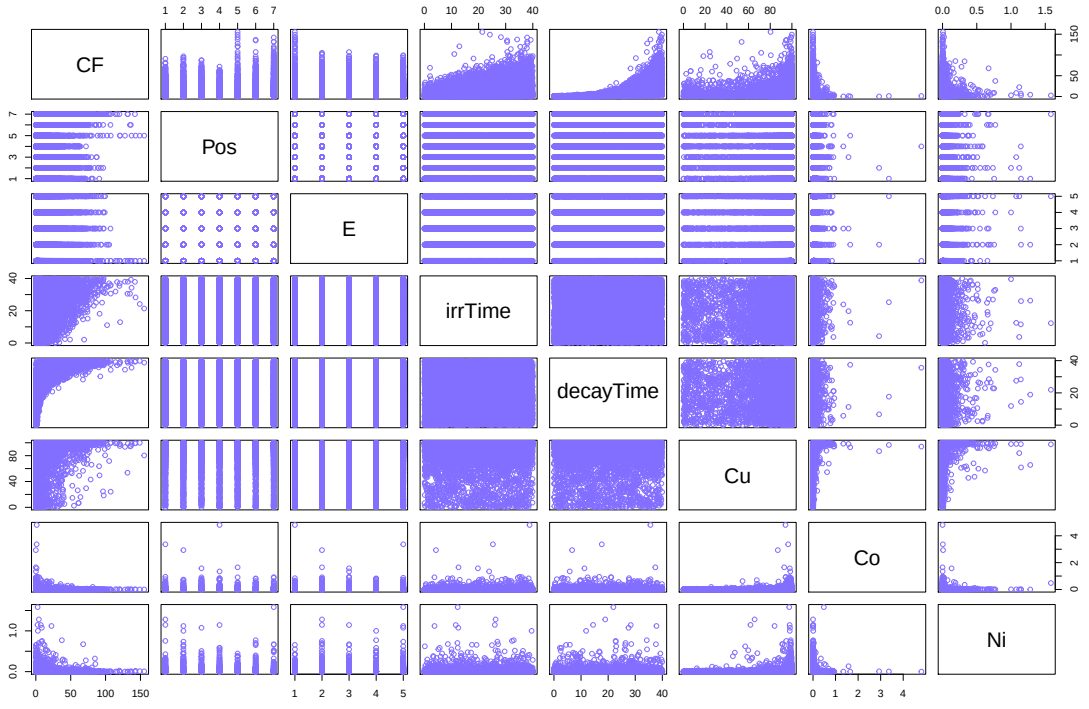


Figure 3.7: Scatter plot matrix of a selected number of variables when 10000 realizations of the multivariate random vector are selected. The CF is calculated for the pairs Ni-63/Co-60
.

From Fig. 3.7 one can see, for instance, that the correlation factor increases with both irradiation and decay times. That is expected because the activity of the Co-60 decays faster than the activity of the Ni-63 and the ratio of the activities Ni-63/Co-60 increases with

time. One can also see that the input space for irradiation and decay times is uniformly sampled. Yet another information can be seen in the window relating the CF with the energy variable. The energy is ordered from smallest (160 MeV) to highest values (7 TeV). It is therefore visible that the correlation factor can have higher values for 160 MeV while it remains fairly constant for all the other energies. This implies that care must be taken in sorting waste that originates from different machines and, in particular, the waste from Linac4 should not be mixed with the waste from all the other accelerators.

Due to the exponential relationship holding between CFs, irradiation and decay times a log-transformation of the response variable increases the quality of the model: the R^2 of the model goes from 0.56 (non-transformed response) to 0.89 (log-transformed reponse). Also other statistics, such as the residual standard error and the F-statistics, become more robust.

A shortened output of the linear model obtained using the statistical language R is showed in Tab. 3.3. The model offers an analytical representation of Eq. 3.27.

Table 3.3: Output of the linear model obtained from 10000 realizations of activated cathodic copper. The pair Ni-63/Co-60 is considered for the calculations of the CF.

Variable	**Estimate**	**Standard error**	**t-value**	**Pr(> |t|)**	**Significance**
14 GeV	-0.051	0.017	-3.032	0.002	**
160 MeV	0.215	0.017	12.778	$< 2 \cdot 10^{-16}$	***
400 GeV	0.081	0.017	4.849	$1.26 \cdot 10^{-6}$	**
7 TeV	0.077	0.017	4.610	$4.07 \cdot 10^{-6}$	***
Position 1	-0.273	0.020	-13.669	$< 2 \cdot 10^{-16}$	***
...	...	...	...	...	...
Irr. Time	0.037	$4.63 \cdot 10^{-4}$	80.711	$< 2 \cdot 10^{-16}$	***
Decay Time	0.124	$4.61 \cdot 10^{-4}$	268.987	$< 2 \cdot 10^{-16}$	***
Cu	-646.7	576.1	-1.123	0.262	
...	...	...	...	...	...

In Tab. 3.3 the estimate represents the coefficient β_i as given in Eq. 3.27. The standard error is the error of the coefficient estimate and the t-value (calculated as the ratio of estimate and standard error) is the parameter used to set the stage for an hypothesis testing. The t-value is the value of the t-statistic for testing wether the corresponding coefficient is different from 0. Finally, the p-value (Pr(> \|t\|)) indicates what would be the probability of obtaining a value such as the one of the t-statistics if the null hypothesis was true. In other words, if the p-value is very low the coefficient is reliable. In Tab. 3.3 this is expressed via

the number of stars in the significance column. A three stars significance implies a good reliability on the given coefficient.

Not all the input variables are useful for prediction and some of them can be neglected when building a model. A number of techniques can be used to select important parameters or the optimum model. In this example we will show a method called "backward stepwise selection" in which a new model is fitted to the data removing one variable at time. The quality of the method can be tested using various statistics such as R^2, adjusted R^2, C_p, Akaike information criterion (AIC) and Bayesian information criterion (BIC). Information on these methods can be found in [Efron and Tibshirani 1993] [James et al. 2017]. A visual output of such a technique is showed in Fig. 3.8.

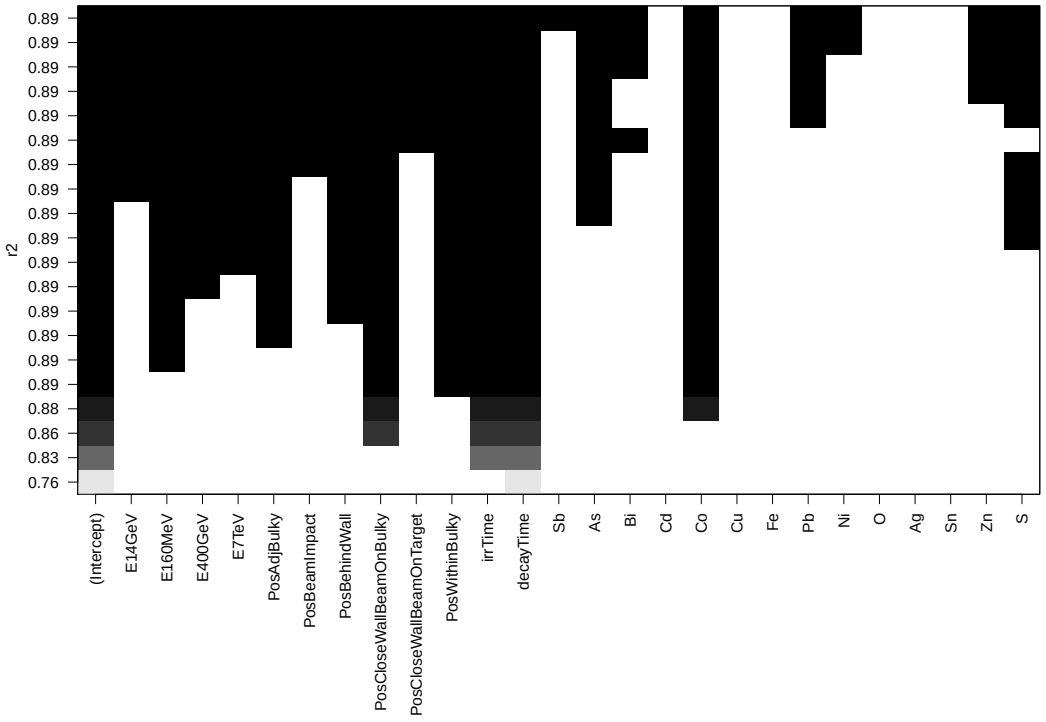


Figure 3.8: Parameter selection using backward stepwise selection coupled with R^2 statistic.

In Fig. 3.8, the plot has the model predictors in the x axis and the value of the R^2 in the y axis. The plot must be read from low to high R^2 and for each R^2 level a new input

parameter is added to the model. For instance, for $R^2 = 0.76$ the feature that allows to best predict the correlation factor is the decay time. Moving upward, if we want to improve the model, we should add the irradiation time ($R^2 = 0.83$). If we want again improve our model we keep adding the features that are in black squares at that R^2 level. It is up to the analyst to decide when to stop in this process. Adding too much variables however can cause overfitting as previously discussed.

For our model, it can be seen that the decay time is a key parameter. This was expected and it is particularly true when the difference between the half-life of DTMs and KN is high, such as in the case studied (the half-lives of Ni-63 and Co-60 are respectively ∼100 and ∼5.3 years). From the plot we can also see that the chemical element Co plays un important role and becomes relevant to predict CFs immediately after decay and irradiation times. This is also physically justified because an increase of cobalt in the chemical composition determinates an increase of Co-60 production with a direct impact on the correlation factor.

We can write the correlation factor as a function of time as follows:

$$CF(t) = \frac{a_{DTM}(t)}{a_{KN}(t)} = \frac{a_{DTM}(0) \times \exp\left(-\frac{\ln(2)}{T_{1/2}^{DTM}} \times \mathrm{t}\right)}{a_{KN}(0) \times \exp\left(-\frac{\ln(2)}{T_{1/2}^{KN}} \times \mathrm{t}\right)} \tag{3.28}$$

where $a_{DTM}(t)$ and $a_{KN}(t)$ are the activities of DTM and KN at the time t, $a_{DTM}(0)$ and $a_{KN}(0)$ are their initial activities and $T_{1/2}^{DTM}$ and $T_{1/2}^{KN}$ are the half-lives of DTM and KN respectively.

It is possible to identify three major cases for Eq. 3.28:

1. $T_{1/2}^{DTM} > T_{1/2}^{KN}$;
2. $T_{1/2}^{DTM} < T_{1/2}^{KN}$;
3. $T_{1/2}^{DTM} \simeq T_{1/2}^{KN}$.

For case 1 we can consider Ni-63 and Co-60, for case 2 Fe-55 and Ti-44 and for case 3 Fe-55 and Na-22. In Eq.3.28 we set the initial activities ($a_{DTM}(0)$ and $a_{KN}(0)$) to 1 for simplicity. We further consider t from 0 up to 30 years.

The evolution of activities and CFs for the cases before mentioned are showed in Fig. 3.9.

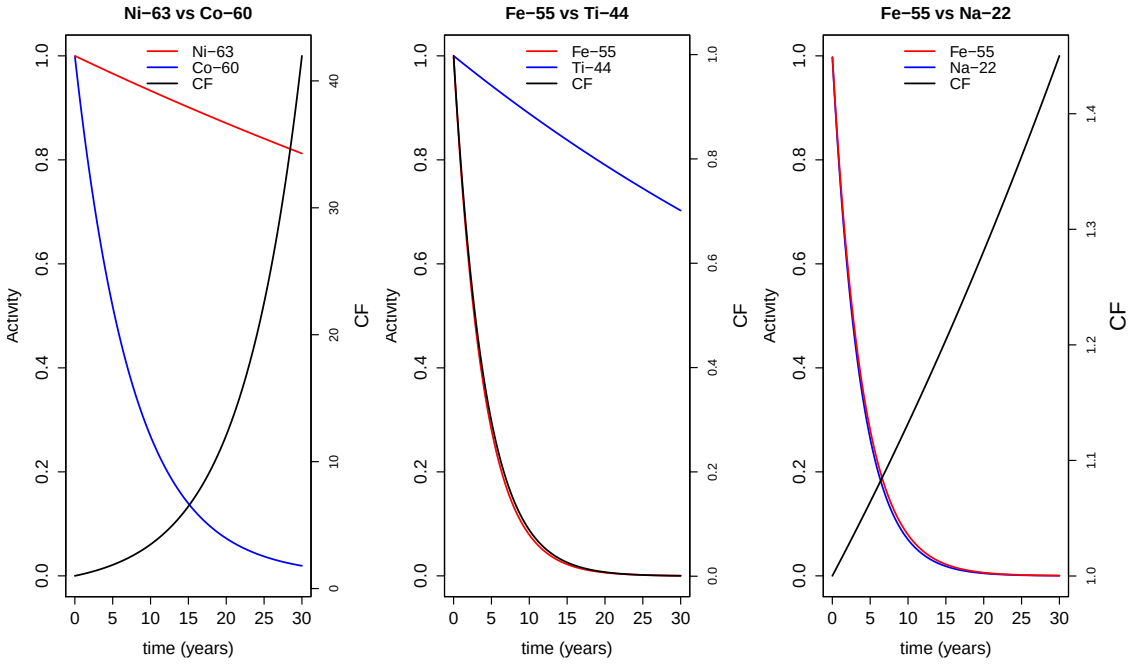


Figure 3.9: Example of evolution of CFs and activity with time. The pairs Ni-63/Co-60 (left plot), Fe-55/Ti-44 (central plot) and Fe-55/Na-22 (right plot) are treated.

When the DTM has an half-life higher than the KN (left plot of Fig. 3.9) the CF increases steadily. For this reason the CF must be corrected with time otherwise the activity of the DTM is underestimated. In the case of Ni-63 and Co-60 the CF can increase over a factor 40 in 30 years.

When the DTM has an half-life lower than the KN (central plot of Fig. 3.9) the CF decreases steadily. The CF must also be corrected with time otherwise the activity of the DTM is overestimated. In the case of Fe-55 and Ti-44 the CF can decrease rapidly towards 0.

Finally, when the DTM has an half-life comparable to that of the KN (right plot of Fig. 3.9) the CF is fairly stationary over a long span of time. The CF does not need to be corrected below a certain threshold which is estimated by the waste producer. In the case of Fe-55 and Na-22 the CF increases of 40% in 30 years.

For the specific case of Ni-63/Co-60 in cathodic copper, we can observe the evolution

of the CF using the simulated activation scenarios. In particular, the scenarios are grouped according to the decay time and a Δt of 5 years.

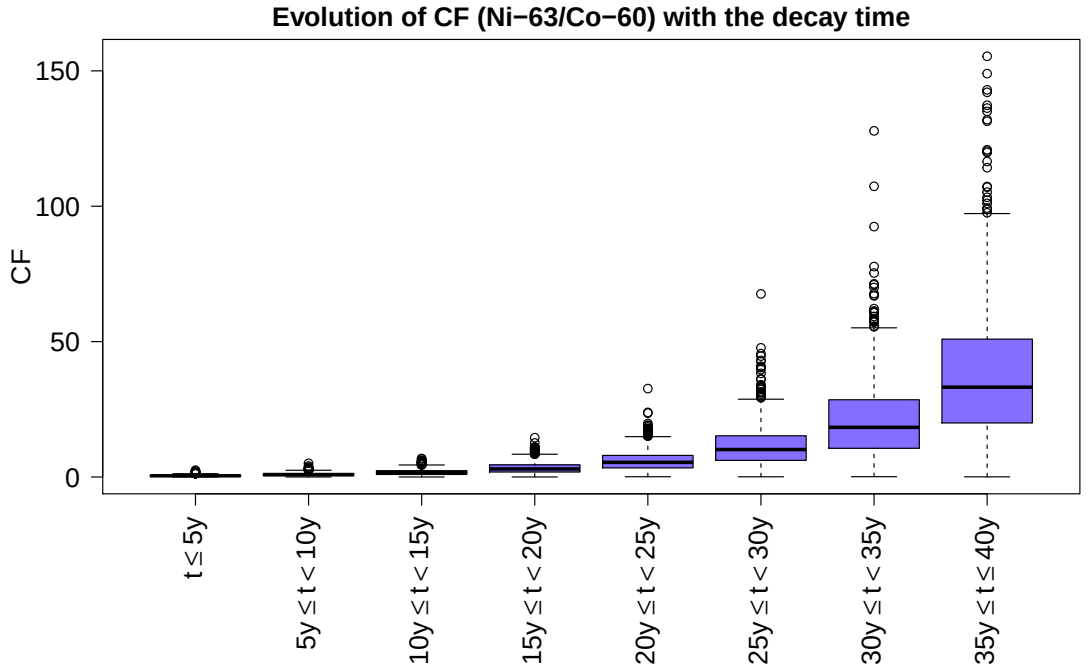


Figure 3.10: Evolution of the CF Ni-63 vs Co-60 with time from 10000 random scenarios.

In Fig. 3.10 one can see the boxplots built using the CFs of Ni-63/Co-60 obtained from 10000 simulated random scenarios. As expected, the correlation factor increases with the decay time. Moreover, the boxplots show the distributions of CFs for each group obtained including the values within a fixed Δt. The outliers are indicators of right skewed distributions.

A number of visual tools are available to evaluate the quality of the fitting in a linear model. Common diagnostic plots are showed in Fig. 3.11.

1. *Residuals vs Fitted*

   The first plot (top-left) indicates non-linearities in the relationship between response and predictors. If a discernible path cannot be found and data are regularly spread around 0, this is an indicator of a good linearity. In practice, the red line should be completely flat. This is the case in our example. The log-transformation of the response linearises the relationship as expected. Few points show however a particular

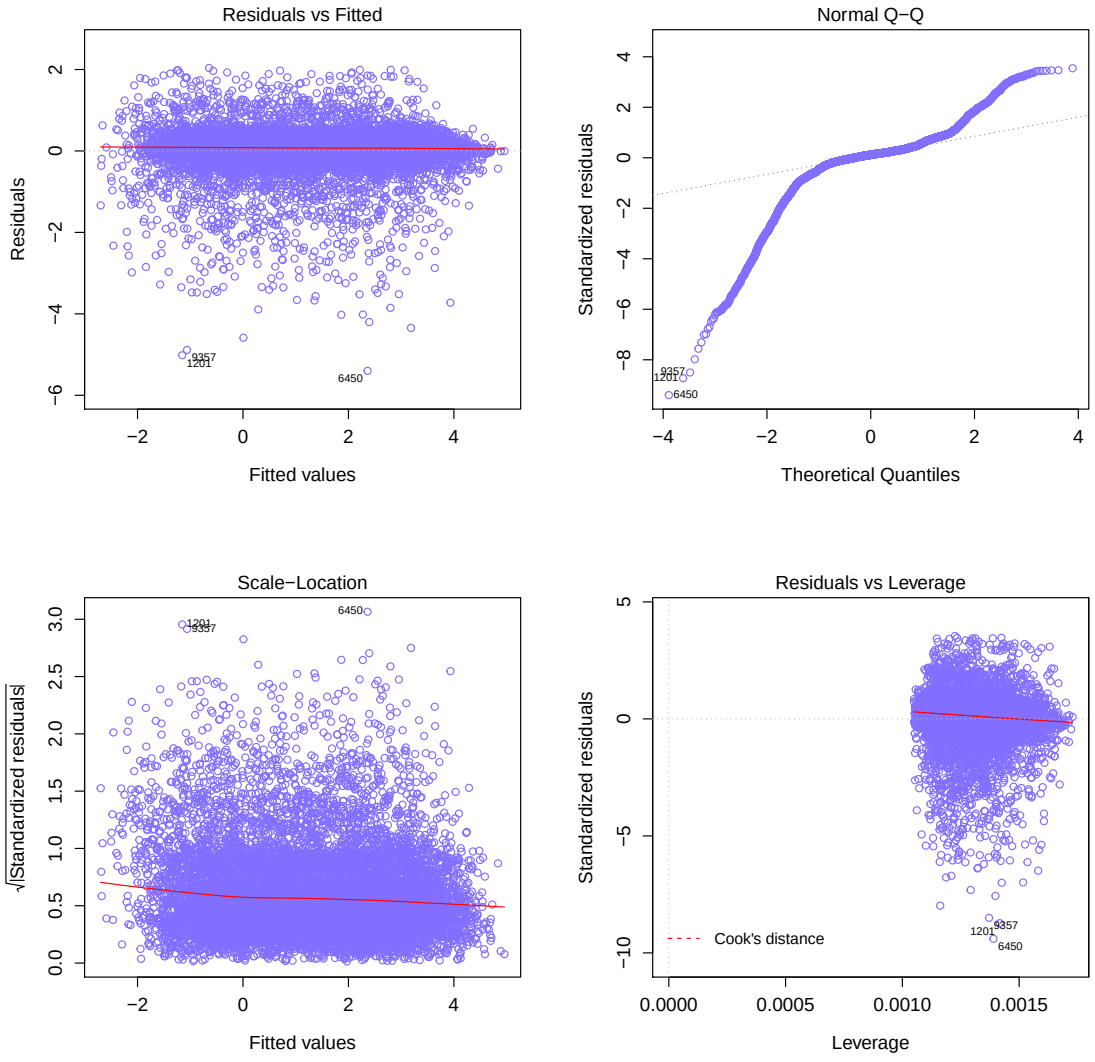


Figure 3.11: Diagnostic plots for the study of the regression model.

high residual and are indicated with the corresponding scenario's number.

2. *Normal qq-plot*

   As a general rule it is expected that residuals follow a normal distribution. This is not the case in our example (see top-right plot of Fig. 3.11). In particular, the tails of our distribution deviate strongly from the Gaussian. Even if a clear pattern cannot be completely demonstrated, the number of non-normal residuals can be considerably reduced if the scenarios with energy equals to 160 MeV are excluded from the model. Another way to visualize extreme values consist of plotting CFs for different energies or position within the tunnel. An example is showed in Fig. 3.12.

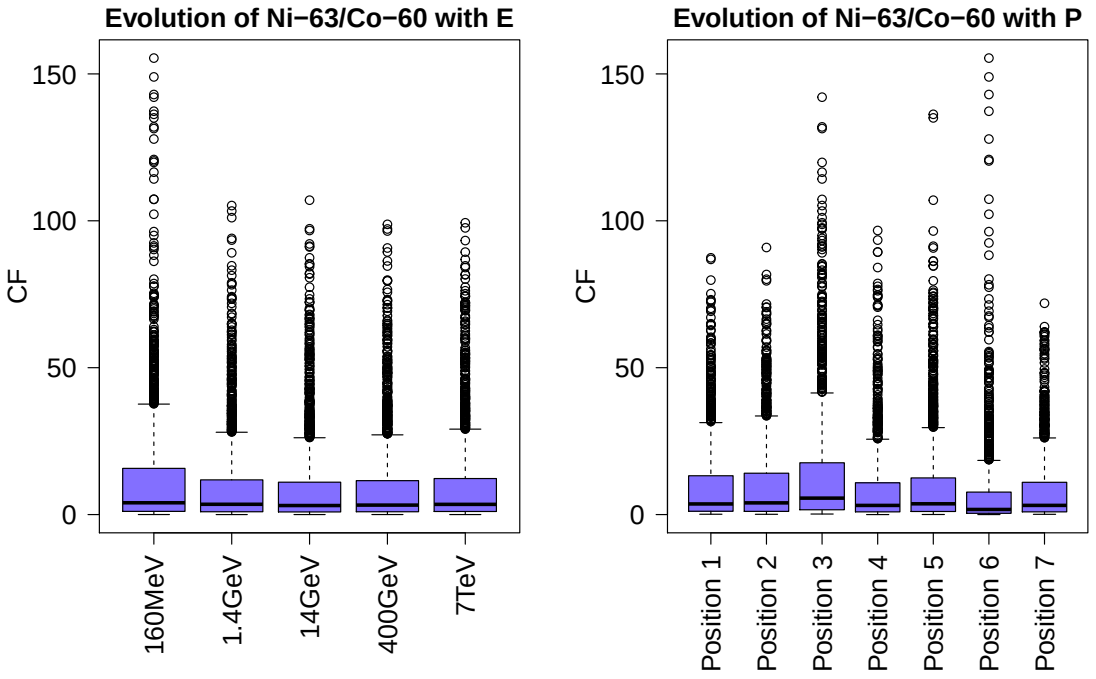


Figure 3.12: Evolution of correlation factors with the energy and the position within the accelerators. The case of Ni-63/Co-60 is showed.

We can see that most extreme values for the correlation factor are produced for energy 160 MeV and Position 6 and 3, corresponding to activations occurring within bulky materials (such as magnets) or close to the wall when the beam interact with the magnets.

3. *Scale-Location*

The scale-location plot (bottom-left of Fig. 3.11) informs about the spread of residuals along the ranges of predictors. It is important for the variance to be fairly constant in the entire predictor interval. In our example this is the case with the exception of few outliers. The statistical term used to indicate an equal variance is "homoscedasticity".

4. *Residuals vs Leverage*

   The last diagnostic plot is used to identify, if any, the influential points. Influential points are outliers that, if included in the model, can affect the model itself. The residual vs leverage plot is given in the bottom-right panel of Fig. 3.11. Influential points can be identified because they lie outside the dotted lines.

To conclude, diagnostic plots are tools that can be employed to analyse the quality of linear models. Even if irregularities are spotted thanks to these tools, the model can still be applied. However the limitation showed by the diagnostic should always be taken into account.

□

3.3.2 Decision trees

In this section we shortly introduce decision trees, which are a powerful alternative to multiple linear regression when estimating correlation factors. The purpose of such a technique is the segmentation of the predictor space into a number of simple regions. A prediction of a given observation is typically based on the mean of the training observations in the region to which the observation belongs.

The optimum segmentation of the input space is obtained via the minimization of the residual standard error, which is the sum of the differences between observations and predictions in each region. This minimization process searches for the single feature, among all the features, that respect such a criterion. Once the feature found, the input space is split in two sub-spaces according to the splitting rule obtained. This technique is commonly called "recursive binary splitting" because it tests each possible input variable in a recursive way and finally splits the input space in two sub-spaces according to the best splitting rule. An example of application follows.

Example 8 *A simplified (pruned) regression tree obtained from 5000 simulated scenarios is showed in Fig. 3.13. The tree is built using correlation factors for Ni-63/Co-60 in cathodic copper.*

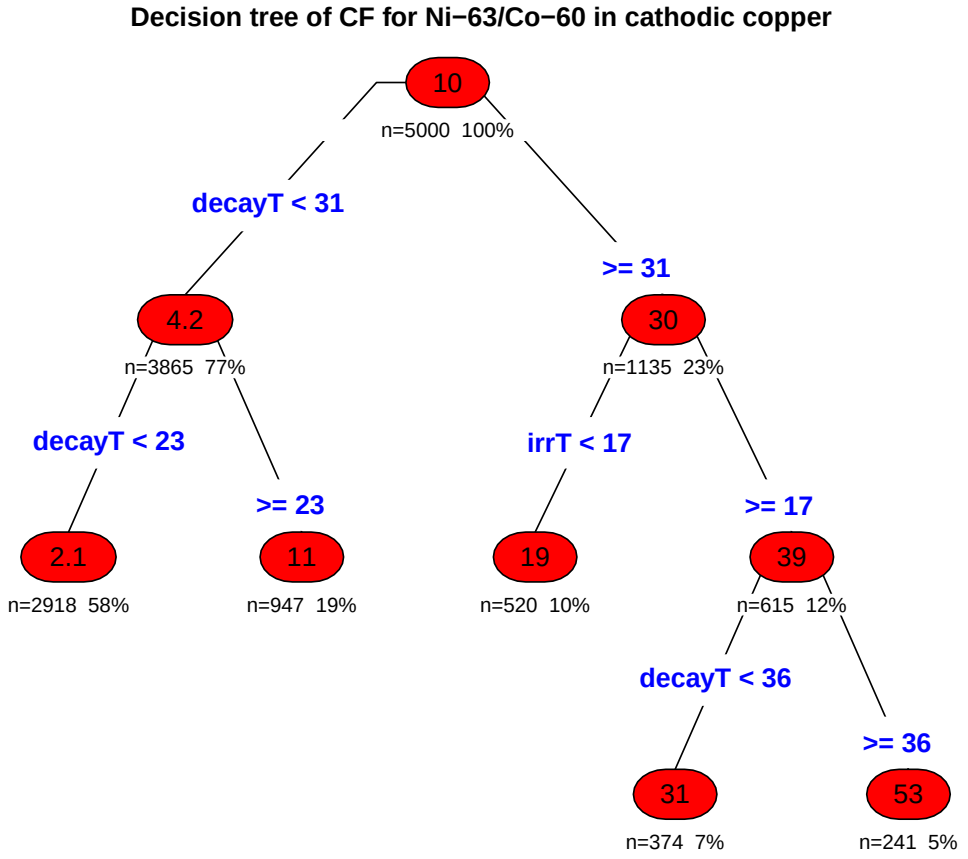


Figure 3.13: Regression tree of the correlation factor Ni-63/Co-60.

Four main information can be obtained from the regression tree of Fig. 3.13. Starting from the top, within the red circle it is given the correlation factor obtained as the average CF calculated from 5000 observations in the training dataset. Below the correlation factor one can see the number of observations ($n = 5000$) and the percentage of observations from the training dataset used to evaluate the CF. Finally, in blue the splitting criterion that minimizes the residual standard error is showed. In this example, the splitting parameter is the decay time (31 years).

Of the 5000 scenarios considered, 77% have a decay time below 31 years (that is 3865

scenarios). The correlation factor Ni-63/Co-60 calculated from the observations with decay time below 31 years is 4.2. The CF in the right branch (decay time above 31 years) is 30. The analysis follows in a top-down simple approach.

From the analysis of the tree we can infer various information. First of all, the decay time plays a major role in determining the correlation factor. The same conclusion was reached when discussing the multiple linear regression. The decay time seems indeed the most important parameter to take into account when we want to evaluate the correlation factors. The irradiation time plays also a strong role but this is less important than the decay time. Finally, the values of the correlation factors are automatically evaluated according to the various splitting criteria that can be used as sorting rules. In fact, if a batch of activated cathodic copper has to be prepared, the best procedure seems to put together waste populations according to their decay time and avoid mixing waste with decay time below and above ∼30 years.

□

A number of techniques can be used to improve the quality of the predictions given by decision trees. One can cite pruning, bagging, random forests and boosting. These techniques are mainly used to increase the robustness of the model or, similarly, to lower the error when making predictions on a test dataset, different from the training dataset.

For instance, "bootstrap aggregation" or "bagging" is a common technique to reduce variance. With this technique, instead of generating a single tree, B trees are grown using repeated bootstrap from the training dataset. A model is obtained from each subset and an average is calculated over the B models to evaluate the statistics of interest.

Another technique used for variance reduction is called "random forest". Random forests are similar to bagging because they lie on a repeated bootstrap. By contrast, only a subset of predictors is however used when searching for splitting criteria. This method basically try to limit the influence of a strong predictor which would be otherwise systematically chosen as first splitting criterion.

Complementary information on these methods can be found in references [Efron and Tibshirani 1993] [James et al. 2017].

Summary

In this chapter we discuss the instrumental and analytical techniques to quantify the activity of the radionuclides in a radioactive waste population.

According to their characteristics, radionuclides can be classified as either ETM or DTM or ITM.

ETM radionuclides can be measured from outside a waste package using non-destructive techniques. At CERN their activity is evaluated using γ-ray spectrometry.

DTM radionuclides are evaluated either via the so-called "scaling factor" method, when an experimental correlation between the activity of DTMs and ETMs can be established, or via the so-called "mean activity method", when radionuclide's activities are uncorrelated and the scaling factor method cannot be applied. The scaling factor method is experimental because a representative number of samples must be collected from a waste population and the activities of both DTMs and a major ETM must be measured in orther to study their correlation.

Finally, the activity of ITM radionuclides is evaluated using the so-called "correlation method". The correlation method lies on extensive Monte Carlo or analytical calculations. Similarly to the scaling factor method, the correlation method looks for correlations between an ETM nuclide and ITMs but on the basis of theoretical activation studies.

In the last section we showed, using statistical learning approaches, that the decay time is the most important parameter to take into account when evaluating correlation and scaling factors.

Chapter 4

Uncertainty quantification

We demand rigidly defined areas of doubt and uncertainty!

Douglas Adams
The Hitchhiker's Guide to the Galaxy.

Introduction

One of the last stages of the radiological characterization process is the quantification of the uncertainty associated with the activity of the radionuclides in a radioactive waste batch (see Fig. 4.1, step G). The complete flow diagram of the characterization process is showed in Ch. 1, Fig. 1.10.

In Ch. 1 we discussed the calculation of the IRAS and its relevance when evaluating the acceptability of a waste population at the disposal facility. We have also seen that national agencies of waste management fix limits for the acceptability of a waste. For the French VLLW disposal pathway these limits were discussed in Sec. 1.3.1.

The calculation of the IRAS depends on the activity of the radionuclides in the waste population and their activity limits. It is important to quantify the uncertainties of these activities to evaluate the distance of the IRAS from the acceptability limits. The quantification of the uncertainty of ETM, DTM and ITM's activities is the topic of the present chapter.

The first section describes the framework of uncertainty quantification with a short

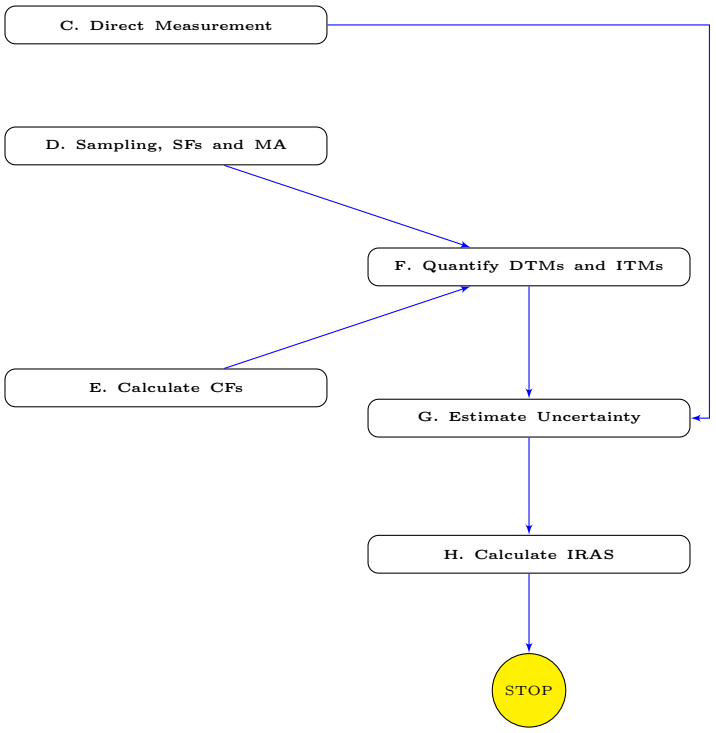


Figure 4.1: Last stages of the radiological characterization process developed at CERN for metallic VLLW. DTMs and ITMs stand for difficult- and impossible-to-measure radionuclides. SF, CF and MA stand for scaling factor, correlation factor and mean activity method.

description of the types of uncertainty. Sections 4.2 to 4.4 discuss the quantification of the uncertainty for ETM, DTM and ITM radionuclides and their contribution to the overall uncertainty budget. Last section introduces the uncertainty propagation associated with the IRAS of a waste package and with the weighted IRAS of a batch of radioactive waste.

4.1 The GUM framework

The expression of a measurement is not complete if only the value of the measurand is expressed. Its associated uncertainty is also required. The uncertainty can be described as a quantitative indication of the quality of the results of a measurement and it is essential to compare measurements or measurement systems between them or to standards but also to draw sound conclusions about test hypothesis.

The Joint Committee for Guides in Metrology[1] (JCGM) reached a consensus about terms and methods to express uncertainty in a large spectrum of scientific applications. The standardization proposed by the working group of the JCGM on uncertainty quantification is described by the known document "Evaluation of measurement data - Guide to the expression of uncertainty in measurement" [JCGM 2008a], often simply referred to as GUM.

In this chapter we will use the terminology and methods introduced by GUM [JCGM 2008a] to estimate the uncertainty of the activities of ETM radionuclides. However, classical uncertainty quantification cannot always be used, especially when some basic hypothesis for its applicability do not hold. In this case, a Monte Carlo based method can be used instead. Two extensions of the GUM where published by JCGM to generalize the methodology to express the uncertainty and they are given in references [JCGM 2008b] [JCGM 2011]. The Monte Carlo based method is described later in this section.

According to [JCGM 2008a] we can identify two classes of uncertainties, called respectively Type A and Type B. Type A uncertainty is evaluated by the statistical analysis of series of observations. By contrast, uncertainty of Type B is evaluated by means that are other than statistical. Common Type B uncertainties are calibration factors and bibliographic values.

The uncertainty of the result of a measurement is, in general, expressed as a standard deviation and, when the result is obtained from the values of a number of other quantities, a combined standard uncertainty is calculated. The rules to combine standard uncertainty are given in [JCGM 2008a][METAS 2012].

When indicating the results of a measurement, the so-called "expanded uncertainty" is given. This quantity defines an interval about the result of a measurement that may be expected to encompass a large fraction of the distribution of values that could be attributed to the measurand. A coverage factor k is commonly used to combine the standard uncertainty to the expanded uncertainty.

A measurand Y, which is the quantity intended to be measured, is often determined

[1] `http://www.iso.org/sites/JCGM/JCGM-introduction.htm`, 30 June 2017.

from N other quantities X_1, X_2, ..., X_N through a function relationship f:

$$Y = f(X_1, X_2, ..., X_N). \tag{4.1}$$

When data is collected we can study the behaviour of Y, through its approximation y. The purpose of uncertainty quantification methods is to quantify or estimate the uncertainty of the dependent variable Y, via the evaluation of the uncertainty on y written as $u(y)$. Similarly, the uncertainties of the various X_i are evaluated from the approximations $u(x_i)$ associated with the data collected upon the independent variables. When the input quantities are uncorrelated, the combined standard uncertainty $u_c(y)$ is the positive square root of the combined variance $u_c^2(y)$:

$$u_c^2(y) = \sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i} \right)^2 u^2(x_i). \tag{4.2}$$

Eq. 4.2 is the first-order Taylor approximation of Eq. 4.1 and it is generally known under the name of "law of propagation of uncertainty". The partial derivative $\partial f / \partial x_i$ is called "sensitivity coefficient" and it is indicated with c_i. Eq. 4.2 is valid only if the linearity hypothesis stands and if the independent variables are uncorrelated. The formulae necessary to take into account non-linearities can be found in [JCGM 2008a].

Let us consider now the case of correlated quantities. In statistics the concept of correlation is used to measure the dependence between events. For instance, when performing γ-spectrometry the temperature T and the humidity H registered in the room can affect the result according to an unknown model. As an example we can imagine that with the variation of T and H both the gross count S_{gross} and background count S_{bkg} are affected in a linear positive way. In this case the two variables S_{gross} and S_{bkg} are said to be correlated. The strength of their correlation can be measured using the so-called "correlation coefficient" r which is given by:

$$r(x_1, x_2) = \frac{s(x_1, x_2)}{s(x_1) \cdot s(x_2)} \tag{4.3}$$

where $s(x_1)$ and $s(x_2)$ are the standard deviation of the variables x_1 and x_2 and $s(x_1, x_2)$,

called "covariance", is given by:

$$s(x_1, x_2) = \frac{1}{n-1} \sum_{i=1}^{n} (x_{1,i} - \overline{x}_1)(x_{2,i} - \overline{x}_2). \tag{4.4}$$

In the previous equations the generic variables x_1 and x_2 were used for generalization.

When input variables are correlated it is necessary to extend Eq. 4.2 to take into account the mixed correlated terms. The combined uncertainty of y is therefore written [JCGM 2008a] [METAS 2012] [JCGM 2011]:

$$u_c^2(y) = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} r(x_i, x_j) u(x_i) u(x_j). \tag{4.5}$$

If the input variables are not correlated ($r(x_i, x_j) \simeq 0$), the mixed terms are cancelled and Eq. 4.5 becomes Eq. 4.2. Let us apply the concept to a simple case.

Example 9 *In a nuclear counting experiment, the net count N_{net} is given by the difference of the gross count N_{gross} and the background count N_{bkg} (which is a counting experience without the radioactive source). We can write this relationship as follows:*

$$N_{net} = N_{gross} - N_{bkg}.$$

If we want to estimate the uncertainty of N_{net} (that we will indicate with y for simplicity) we can use Eq. 4.5. In the calculations that follow, N_{gross} is indicated by x_1 and N_{bkg} is indicated by x_2 ($y = x_1 - x_2$).

$$\begin{aligned} u_c^2(y) = & \frac{\partial f}{\partial x_1} \cdot \frac{\partial f}{\partial x_1} \cdot r(x_1, x_1) \cdot u(x_1) \cdot u(x_1) + \\ & + \frac{\partial f}{\partial x_1} \cdot \frac{\partial f}{\partial x_2} \cdot r(x_1, x_2) \cdot u(x_1) \cdot u(x_2) + \\ & + \frac{\partial f}{\partial x_2} \cdot \frac{\partial f}{\partial x_1} \cdot r(x_2, x_1) \cdot u(x_2) \cdot u(x_1) + \\ & + \frac{\partial f}{\partial x_2} \cdot \frac{\partial f}{\partial x_2} \cdot r(x_2, x_2) \cdot u(x_2) \cdot u(x_2). \end{aligned}$$

After simple algebraic manipulations, the evaluation of the partial derivatives and considering that $r(x_i, x_i) = 1$ and that $r(x_i, x_j) = r(x_j, x_i)$ we can write:

$$
\begin{aligned}
u_c^2(y) = &\left(\frac{\partial f}{\partial x_1}\right)^2 \cdot u^2(x_1) + \\
&+ \left(\frac{\partial f}{\partial x_2}\right)^2 \cdot u^2(x_2) + \\
&+ 2 \cdot \frac{\partial f}{\partial x_1} \cdot \frac{\partial f}{\partial x_2} \cdot r(x_1, x_2) \cdot u(x_1) \cdot u(x_2) = \\
&= u^2(x_1) + u^2(x_2) - 2 \cdot r(x_1, x_2) \cdot u(x_1) \cdot u(x_2).
\end{aligned}
$$

Replacing the counting symbols to y, x_1 and x_2 we obtain the combined uncertainty of the net count $u_c^2(N_{net})$:

$$u_c^2(N_{net}) = u^2(N_{gross}) + u^2(N_{bkg}) - 2 \cdot r(N_{gross}, N_{bkg}) \cdot u(N_{gross}) \cdot u(N_{bkg}).$$

□

Finally, the expanded uncertainty U associated with y, can be expressed as the product of the combined uncertainty and the coverage factor k:

$$U = k \times u_c(y). \tag{4.6}$$

In practice the coverage factor is often chosen to be 2 or 3 and the best estimate of the measurand Y is reported as $Y = y \pm U$. The value of the coverage factor is however linked to the underlying distributions of the input variables and, theoretically, it should be demonstrated that either the input variables are distributed according to a normal distribution or that the variances of each input variable are approximately equal. In reality it can be demonstrated that for samples of small size (i.e. <30) the t-distribution is a better approximation to describe the result of a measurement reported as the mean of repeated observations (uncertainty of Type A) [METAS 2012]. The values of the coverage factor k can therefore be obtained from a t-table.

In some cases, the GUM method cannot be used for the calculation of uncertainties. Some of the basic hypotheses for the GUM method to be applicable are the "almost" linearity of the function f, the validity of the central limit theorem requiring that a normal

distribution or a t distribution could be associated with the values of the measurand Y and that the input variables X_i should not be correlated if their degree of freedom is not infinite [METAS 2012]. Monte Carlo methods can be used as an alternative to estimate the uncertainty in such cases.

The Monte Carlo scheme is based on some assumptions made about the probability distributions of the input variables $g(X_i)$ and the functional relationship f between input and output variables. A simplified scheme of the Monte Carlo method with three input variables is given in Fig. 4.2.

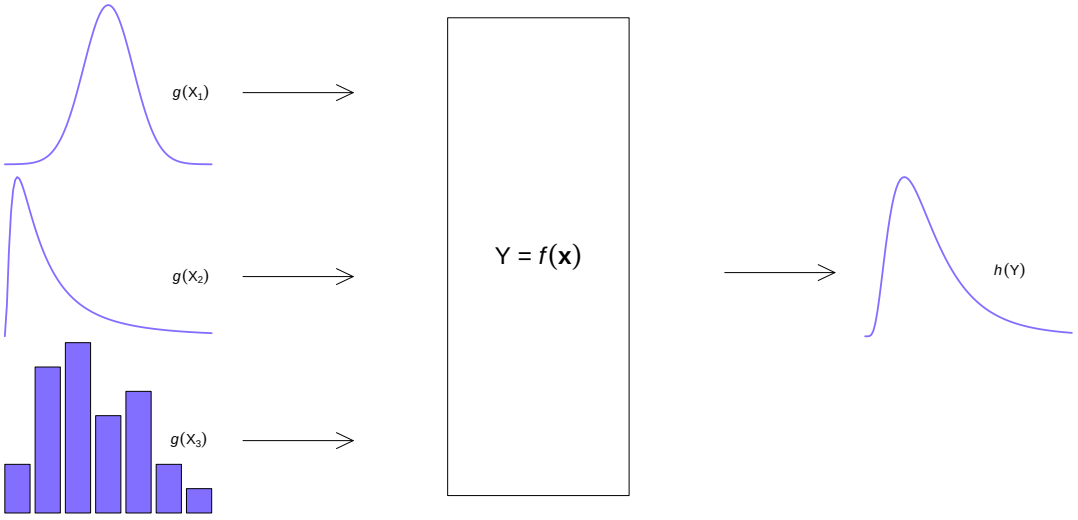


Figure 4.2: Monte Carlo scheme used to estimate the uncertainty of a measurand Y based on the distribution of three input variables via a functional relationship f. Adapted from [JCGM 2008b] [METAS 2012].

The main steps of the Monte Carlo methodology are the following [JCGM 2008b] [METAS 2012]:

1. The measurand Y is defined;

2. The input variables X_i are determined;

3. The functional relationship between Y and X_i is expressed mathematically as $Y =$

$f(X_1, X_2, ..., X_n)$;

4. A distribution $g(X_i)$ is associated with each entry variable X_i; this distribution is obtained from the available information;

5. The probability density distribution $h(Y)$ is calculated propagating the distributions $g(X_i)$ through the function f. This is done via the random extraction of n values of $\mathbf{x} = (X_1, X_2, ..., X_n)$ according to the joint distribution $g(\mathbf{x})$ and the injection within the function f to obtain a realization of Y. This step is performed M large number of times (we commonly choose $M = 10^4$). At the end we obtain a list, of dimension M, of values y representing the measurand $Y : y = \{y_1, y_2, ..., y_M\}$;

6. The best estimator of $h(Y)$ is the empirical distribution $h(y)$;

7. The uncertainty $u(y)$ is calculated as the empirical standard deviation of $h(y)$;

8. The extended uncertainty is finally calculated according to the coverage factor k chosen.

In the following sections we use both classical and Monte Carlo techniques to quantify the overall uncertainty of the IRAS.

4.2 Quantification of ETM's uncertainty

In this section we discuss the principal terms contributing to the uncertainty of ETM's activity. Within the radiological characterization framework of VLLW produced at CERN, ETM radionuclides occupy a leading role because, as a general rule, they are the major contributors to the IRAS. This is mainly due to their low activity limit as discussed in Sec. 1.3.1. The uncertainty of the activity for this family must therefore be quantified accurately because its impact on the uncertainty of the IRAS is preponderant.

We start the discussion with the analytical uncertainties (Type A). In the second part we introduce some techniques that can be used to estimate the uncertainty associated with the bias term. The bias is linked to activity and weight heterogeneities but also to the measurement choices made at CERN.

4.2.1 Analytical uncertainty

In general, γ-spectrometries are performed systematically for each waste package (census). We discuss in this section the parameters that must be taken into account to quantify the analytical uncertainty (Type A) of the activity of a package of radioactive metallic waste measured via in-situ γ-spectrometry.

In Ch. 3, Sec. 3.1 we discussed in detail the formula to calculate the specific activity of γ-emitters when performing γ-ray spectrometry. If the activity a_{ETM} of an ETM radionuclide is calculated from a single peak without interferences and the correction terms for irradiation and decay are neglected, its analytical combined uncertainty $u_c(a_{ETM})$ can be calculated from [CANBERRA 2009]:

$$\frac{u_c(a_{ETM})}{a_{ETM}} = \sqrt{\left[\frac{u(S_{net})}{S_{net}}\right]^2 + \left[\frac{u(m)}{m}\right]^2 + \left[\frac{u(\epsilon)}{\epsilon}\right]^2 + \left[\frac{u(I_\gamma)}{I_\gamma}\right]^2} \tag{4.7}$$

where $u(S_{net})$ is the uncertainty of the net peak area, $u(m)$ is the uncertainty of the weight of a waste package or sample, $u(I_\gamma)$ is the uncertainty of the γ emission probability and $u(\epsilon)$ is the uncertainty of the efficiency calibration given by:

$$\frac{u(\epsilon)}{\epsilon} = \sqrt{\left[\frac{u(\epsilon^{'})}{\epsilon^{'}}\right]^2 + [\rho l \cdot u(\mu(E))]^2 + [\mu(E) \cdot u(\rho l)]^2} \tag{4.8}$$

with $u(\epsilon^{'})$ the uncertainty of the uncorrected efficiency, $u(\rho l)$ the uncertainty of the average sample mass per unit area and $u(\mu(E))$ the uncertainty of the attenuation coefficient.

Eq. 4.7 is a simplified version of the formula used to quantify the uncertainty of ETMs. Complementary random and systematic errors can for instance be added by the analyst to take into account other phenomena such as heterogeneity. More complex cases, including interference correction and multiplets, require an advanced numerical treatment. Their discussion is outside the scope of the present thesis and their treatment can be found in reference [Lépy et al. 2015][CANBERRA 2009].

The formulae given before are also ideal in the sense that they assume a specific activity distribution within a waste package (generally uniform), a uniform distribution of the weight within the package and also a perfect knowledge of the angular response of the

detector. In fact the first two terms generate a bias and the third term (detector response) is affected by an uncertainty that is quantified by the detector producer using Monte Carlo techniques. In practice, extra uncertainties are added to the efficiency to take into account the real characteristics of the detector. The uncertainty of ϵ decreases with the energy and spans between $\sim$15% at 50 keV and 2-3% at 2 MeV.

When performing γ-spectrometry measurements on large objects, such as the radioactive waste packages used at CERN (see Ch. 3, Fig. 3.4) one of the major difficulties is the establishment of activity and mass distributions within the package. These distributions have a direct impact on both the efficiency calibration and, consequently, the uncertainty of the specific activity of the ETMs. We discuss these bias terms in the next section.

4.2.2 Bias

Heterogeneity of both activity and weight has a direct impact on the evaluation of the activity because of its effect on the efficiency calibration $\epsilon(E)$. The standard software used at CERN to build efficiency calibration functions [CANBERRA 2009] assumes that the activity within a sample is distributed either uniformly or according to a predefined distribution. The real distribution of the activity is however unknown and cannot be estimated easily. From a statistical point of view this physical phenomenon represents a bias. Weight distributions are also considered uniform and this assumption has a direct impact on the terms $\mu(E)$ and ρl of Eq. 4.8.

To limit the impact of such biases a number of operational mitigating actions are put into place at CERN. These actions are described below.

- The selection of the waste items to put into a given package is made according to the best knowledge of the radiological history of the waste. As far as possible, items originating from different sites and with different history are not mixed.
- Since the dose rate of an item is proportional to its γ activity, dose rate measurements are used to screen out hot-spots or extreme values. Only low heterogeneous waste with respect to their dose rate are combined within a waste package.
- The vast majority of waste items are cut, mixed and compacted for operational

and technical reasons. These operations are useful to decrease activity and weight heterogeneities. At the end of these operations an apparent density is calculated for the waste packages. This density is useful when determining the calibration function. The density is called apparent because it does not reflect the material density but only the volume reduction factor of the items conditioned within a waste package. For instance, if common density for different grades of copper is in the range 8.5 up to 9 g/cm^3, apparent density of CERN waste packages filled with copper waste spans in the range 1-2.5 g/cm^3. This is mainly do to our present compaction capacity. It is estimated that after the process of cutting, mixing, compacting and packaging we can operate a volume reduction up to a factor 4.5 [Michaud and Zaffora 2013].

- Multiple measurements, from different sides of the waste package, are performed. This action allows us to evaluate the average activity of ETMs in a waste package in a more robust way. Even if heterogeneities are present, their impact on the average activity is mitigated by performing multiple measurements from different angles.

- Extreme cases were tested considering the activity concentrated in particularly unfavourable positions within a waste package. For instance, we simulated very active point sources at the centre of the waste package and evaluated the efficiency calibration function in such a case. When this assumption is made we can obtain specific activities which are orders of magnitude (1-to-3) above the ones obtained for an uniform distribution [Nagy and Zaffora 2016]. This test however has to be considered as unrealistic for practical applications. In fact, the preliminary dose rate screening of each item forbids to choose items with extreme dose rates. Such items are not conditioned within standard packages and are treated as special cases.

The choice of a uniform activity distribution will end up with a bias that underestimates or overestimates the true activity of the ETMs in a waste package. Over large populations however such biases will eventually cancel out when we estimate the final average activity of the waste population. In order to estimate the uncertainty generated by the weight and activity biases (often referred to as systematic uncertainty) we performed a test on 56 containers of legacy metallic waste consisting of compacted blocs.

We performed 224 γ-spectrometry measurements, 4 per package, on 4 sides of the containers. We quantified the average activity of the key nuclide Co-60 for each container taking into account the uncertainty associated with each measurement. The average activity $\overline{a}(RN)$ of a generic radionuclide RN can be calculated as follows:

$$\overline{a}(RN) = \frac{\sum_i \frac{a_i(RN)}{u_i^2(a(RN))}}{\sum_i \frac{1}{u_i^2(a(RN))}} \tag{4.9}$$

where $u_i^2(a(RN))$ is the square of the absolute uncertainty of the activity of RN from the measurement performed on the i^{th} side of the waste package.

We calculated the relative error ϵ_{rel}^i for the activity measurement of Co-60 on a generic side of the waste package as follows:

$$\epsilon_{rel}^i = \frac{|a_i(\text{Co-60}) - \overline{a}(\text{Co-60})|}{\overline{a}(\text{Co-60})}. \tag{4.10}$$

Using this process we built a distribution (right-skewed) of relative errors. We found that the median error, due to weight and activity heterogeneities, is 23.3%. Even if we cannot distinguish the impact of weight and activity distribution on the activity of radionuclides in a radioactive waste package we can use this study to have a first approximation of the bias extent. Such a study is however valid only for metallic waste conditioned within the standard waste packages previously described (see Fig. 3.4, Ch. 3).

At CERN the estimation of the average ETM's activity in the past years was made performing 4 γ-spectrometries on the 4 sides of the standard containers. In 2015 we performed a study to demonstrate the impact of the reduction of γ-spectrometry measurements (from 4 to 2) on the IRAS. The results are described on the report of reference [Magistris and Zaffora 2015]. This test was performed on the same 56 radioactive waste packages containing compacted metallic blocks previously described and disposed of as VLLW waste in 2013-2014 [Michaud and Zaffora 2013]. We found that the reduction of the number of measurements has an impact below 1% on the IRAS of a package if the two measurements are performed on the long sides of the waste package. The maximum variation of the IRAS was 0.88%. The specific activity of a single γ-emitter can however vary highly, with respect

to the average, when the number of measurements is reduced due to the weight/activity heterogeneities within a waste package.

To estimate the impact of such a new measurement procedure on the average activity of a waste package we calculated, for each waste package, the average activity obtained from two sides. For each waste package 6 combinations of sides are possible. We found that the median relative uncertainty associated with the reduction of the number of measurements is 13.2% if the average activity obtained from 4 measurements is considered as the reference.

The estimations of the two biases can be summed quadratically to quantify the total bias contribution of the Mean Squared Error (see Eq. 3.15, Sec. 3.2.3).

4.3 Quantification of DTM's uncertainty

In Ch. 3 we discussed the scaling factor and mean activity methods to quantify the activity of DTM radionulides. In Sec. 3.2.3 we have also seen which elements most contribute to the random variability of a statistics $\hat{\theta}$ when sampling a waste population. In particular, we have seen that the terms that most contribute to the random uncertainty are the sampling and the analytical variabilities.

As a general rule, the analytical uncertainty is often low and fairly constant since the measurements on samples are performed within a rigid and standardized measurement framework. For 1 Bq/g DTM's activity it is common to achieve analytical uncertainties of 5% for a coverage factor $k = 1$.

In the present section we discuss the techniques that can be used to quantify the uncertainty $u(a_{DTM})$ of the DTM radionuclides based on the estimation of each term of the random uncertainty. The discussion on the bias term follows.

4.3.1 Analytical uncertainty of scaling factors

The activity of the generic DTM radionuclide on a waste package is calculated using Eq. 3.7. The activity is the result of the multiplication of the SF and the measured activity of the KN in that package. We have seen that the value of SF can be estimated either as the geometric mean of a series of ratios or using a linear model. If the geometric mean SF

is used (see Sec. 3.2) we can obtain the combined uncertainty of the DTM activity from Eq. 4.5:

$$\frac{u_c^2(a_{DTM})}{a_{DTM}^2} = \frac{u^2(\overline{G}_{SF})}{\overline{G}_{SF}^2} + \frac{u^2(a_{KN})}{a_{KN}^2} + \frac{2 \cdot r(\overline{G}_{SF}, a_{KN}) \cdot u(\overline{G}_{SF}) \cdot u(a_{KN})}{\overline{G}_{SF} \cdot a_{KN}} \tag{4.11}$$

where the term $u(\overline{G}_{SF})$ can be estimated by the standard error of the geometric mean [Harding et al. 2014][Zaffora et al. 2017b]:

$$u(\overline{G}_{SF}) = \overline{G}_{SF} \times \frac{D}{\sqrt{(n-1)}} \tag{4.12}$$

where $\overline{G}_{SF}$ and D are given by Eqs. 3.10 and 3.12.

In Eq. 4.11 the term $u^2(a_{KN})$ is the variance associated with the measured activity of the KN and evaluated according to the methods discussed in Sec. 4.2. The correlation between a constant (SF) and a random variable (a_{KN}) being 0, the mixed term is neglected. We can finally write:

$$\frac{u_c(a_{DTM})}{a_{DTM}} = \sqrt{\frac{u^2(\overline{G}_{SF})}{\overline{G}_{SF}^2} + \frac{u^2(a_{KN})}{a_{KN}^2}}. \tag{4.13}$$

If we use linear models, the uncertainty $u(a_{DTM})$ can be evaluated using the confidence interval around the regression line. In this case we can write [Olive 2007]:

$$u_c(a_{DTM}) = s(a_{DTM}) \cdot \sqrt{\frac{1}{n} + \frac{(a_{KN} - \overline{a}_{KN})^2}{(n-1) \cdot s^2(a_{KN})}} \tag{4.14}$$

where $s(a_{DTM})$ is the sample standard deviation of a_{DTM}, n is the number of samples collected, $\overline{a}_{KN}$ is the average KN activity and $s^2(a_{KN})$ is the variance of the KN activity evaluated from the sample.

4.3.2 Analytical uncertainty of the mean activity

We have seen in Sec. 3.2.2 that if the scaling factor method cannot be applied we can simply calculate the average DTM activity and apply this value to each package of the waste population.

The sampling variability can be easily obtained from the sampling variance according to the sampling method adopted. In Ch. 3, Sec. 3.2 we discussed a number of sampling methods including simple random, systematic and stratified sampling. For each one of them the sampling variance can be estimated using classical formulae given in a number of references, such as in [Cochran 1977][Gilbert 1987] and [Myers 1997]. This variance corresponds to the sampling variability of the random uncertainty. It is important to recall that, if non-probabilistic approaches are used to collect the sample, a variance cannot be properly estimated and it can be difficult to quantify the impact of the bias.

The analytical variability is taken into account including the uncertainty from the measurement system into the propagation of uncertainties. This can be done using either a weighted mean or the bootstrap standard error as discussed in Sec. 3.2.2. The weighted mean can be calculated using a formula of the form of Eq. 4.9.

4.3.3 Bias

A number of biases can occur when estimating the activity of DTM radionuclides. We can name the presence of activity heterogeneities in the waste population, a non-probabilistic sampling, the choice of the wrong KN and an incomplete prediction of the radionuclides in the waste population [IAEA 2007].

In order to quantify the bias that can occur as a consequence of activity heterogeneity we performed an experiment over a waste population of shredded copper. The findings of the study can be found in [Zaffora et al. 2016]. We collected a census sample from the waste population and measured the activity of Co-60 (KN), Ni-63, H-3 and Fe-55.

We compared the results in terms of population activity when different sampling strategy were applied. In particular, we considered simple random and systematic sampling and the bootstrap. Comparing the results of each technique with the activity from the census we estimated the strength of each sampling method.

As discussed in Ch. 3, Sec. 3.2.2, the bootstrap offers the possibility to evaluate the bias when estimating a statistics of a population (such as the mean activity). We tested the bootstrap for a different number of cases going from the selection of 5 samples out of

87 and 250 repetitions up to 20 samples and 1000 repetitions.

During the campaigns performed to collected samples we rarely choose more than 5% of the total units of a waste population. Using this value as a reference and for 1000 repetitions, we can estimate the bias affecting the average amount of Ni-63 in the waste population. We found that the relative error (when values below the limit of detection are not used) is 21.2%.

In absence of more precise estimations this amount will be considered as an estimator of the extent of the bias for the evaluation of the uncertainty of DTM's activity.

The bias contribution due to non-probabilistic sampling is difficult to estimate and often cannot be quantified. It depends strongly on the knowledge of the waste population and the sampling strategies adopted by the waste producer. A solution often adopted at CERN to mitigate its impact consists of choosing appropriate high-dose samples. Biased sampling was described in Ch. 3, Sec. 3.2.

The bias associated with the choice of the wrong KN has a limited impact for the characterization of legacy radioactive waste at CERN. This is due to the fact that the majority of radionuclides with short half-lives have decayed at the time of the characterization. Only limited key nuclides therefore have the properties to be selected and were discussed in Sec. 3.1.

Finally, the bias associated with the incomplete prediction of radionuclides is common for ETM, DTM and ITM radionuclides. We already discussed the solution adopted for ETM radionuclides. Complementary information is given in the next section.

4.4 Quantification of ITM's uncertainty

The quantification of the uncertainty associated with the activity of ITM radionuclides comes from the theoretical activation studies used to evaluate correlation factors (see Sec. 3.3) and from the measure of the activity of the KN. The uncertainty quantification of the activity of the KN was already discussed in Sec. 4.2. We will focus here on the uncertainty of the activation studies performed with Actiwiz [Theis and Vincke 2012].

When dealing with legacy waste we have seen that common information such as the

exact elemental composition or the radiological history of a waste is unknown or incomplete. We defined a mixed multivariate random vector **S** to generate a large number of realizations in order to take into account the potential variability of the input factors.

From a high number of realizations of the vector **S** we can build a distribution for the CF and calculate its dispersion. From a formal point of view, the random extraction of input parameters and the construction of a distribution for the correlation factor reflect the logic of a Monte Carlo method. The structure of this method was described in Sec. 4.1. The distribution of the output variable, such as the distribution of CF, is however affected by a number of limitations that we would like to describe in this section.

When we build distributions for the input parameters, we are making hypotheses which rely on the knowledge of the radioactive waste population. For instance, we decide to include (or not) a specific chemical element in the generic composition of a material. Presently, the approach used is conservative in the sense that we add all the potential trace elements or contaminants in order to avoid that some production channel is missed. This approach however can predict a list of radionuclides that are not produced because, in reality, the father element was absent from the elemental composition.

Of course, the opposite can also be true and we miss the production of some radionuclides because we decided not to include a specific chemical element. This second option has however a limited impact in the radiological characterization of waste for two reasons. First of all we use conservative elemental compositions with a long list of chemical elements by default. Secondly, we perform systematic γ-spectrometries on each single waste package. If a radionuclide that was not predicted is found by measurements, the elemental composition adopted is questioned and optimized according to the measurement results.

The most unfavourable case would be that only radionuclides that do not emit γ-photons are produced and not measured. Measurements of γ-rays would not give indications of bias in the chemical composition. Also this case is unrealistic because a long list of activation studies has been performed at particle accelerators during last years (see for instance references [Mitaroff and Silari 2002] and [Brugger et al. 2006]) and a good knowledge was gained about the radionuclide inventories obtained from activation at particle accelerators. The tools presently used at CERN, including Fluka [Battistoni et al. 2006][Ferrari

et al. 2005] and Actiwiz [Theis and Vincke 2012], are robust for radionuclide inventory predictions.

A second typology of bias comes from the random extraction of input parameters, such as the beam energy and the position within the accelerator's tunnel. Uniform random extraction of energy beam implies that each energy is extracted about 20% of the time. Of course, a batch of waste cannot systematically come from multiple accelerators. In Ch. 3 we have seen however that the correlation factors are comparable for all energies, exception made for the Linac 4 (which generates higher CFs). Extracting from all the available energies is therefore acceptable, exception made for the Linac 4 which has the tendency to push the correlation factor towards higher values. When the Linac 4 is included on the simulations we are systematically overestimating the average value of the CF.

In terms of irradiation positions we also make the hypothesis of a random extraction from a uniform distribution for the 7 standard levels of the variable. This choice is made according to the present capabilities of the second version of Actiwiz [Theis and Vincke 2012] which allows us to choose only fixed positions. This limitation will be overcome by the future version 3 of Actiwiz which will be able to accept specific particle fluence spectra according to the need of the analyst. We have however discussed in Sec. 2.1.3 that the positions presently available represent fairly well the most common CERN activation scenarios for hadron machines. The capabilities of the new Actiwiz version will also accept spectra of electron machines, such as the LEP.

The random extraction from a discrete uniform distribution attributes the same weight to the 7 positions. This is in fact a bias term because the real waste do not necessarily originates from a perfect mix of materials installed over all the irradiation positions. This approach however is considered as conservative because it oversamples some positions within the tunnel (such as the beam impact area) in which only a small fraction of the materials can be installed.

Finally, the irradiation and decay times have a high impact on the value of the correlation factor. In absence of precise information about the radiological history of a waste population we can only attribute uniform distributions to these two predictors. The date of reception at the storage can be used as a filter for more realistic decay times.

The mixed multivariate random vector **S**, introduced in Ch. 2, is a useful tool to predict a complete radionuclide inventory and to build distributions for the theoretical correlation factors. The correlation factors so calculated represent a first estimation of the scaling factors for the radionuclides that have a realistic impact on the IRAS. The radionuclides left over (ITM) often have a negligible impact on the IRAS but are included in the radionuclide inventory that is declared to the national waste management agencies.

As previously discussed in Ch. 3, Sec. 3.3, the geometric correlation factor is a good central tendency estimator for the distributions of CFs. Both heterogeneity and sampling error can be estimated from these distributions. It is however important to remember that the particle fluence spectra are obtained from Monte Carlo simulations in which some of the data, such as cross-sections and their uncertainties, is obtained from standard libraries. The uncertainties associated with these parameters should be considered as Type B.

At the time of writing (June 2017) we consider that the standard error of the geometric correlation factor reflects alone the uncertainty and bias associated with the correlation factor. This uncertainty can be evaluated in a similar way as for the scaling factors. In particular, if we indicate with $\overline{G}_{CF}$ the geometric correlation factor and with $u(\overline{G}_{CF})$ its uncertainty, we can write the uncertainty of the ITM's activity as follows:

$$\frac{u_c(a_{ITM})}{a_{ITM}} = \sqrt{\frac{u^2(\overline{G}_{CF})}{\overline{G}_{CF}^2} + \frac{u^2(a_{KN})}{a_{KN}^2}}. \tag{4.15}$$

4.5 Propagating uncertainties

The purpose of the present section is to introduce a formula to quantify the uncertainty of the IRAS. After the IRAS is calculated (step H of Fig. 4.1) its uncertainty must be quantified to evaluate the distance that separates IRAS from the acceptability limits.

Using the uncertainty propagation formula (Eq. 4.5) we can write the combined uncer-

tainty of the IRAS $u_c(IRAS)$ as follows:

$$u_c^2(IRAS) = \sum_{i=1}^{n} \frac{u^2(a_i)}{AL_i^2} + \sum_{\substack{i=1 \\ i \neq j}}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} \frac{r(a_i, a_j) \cdot u(a_i) \cdot u(a_j)}{AL_i \cdot AL_j} \tag{4.16}$$

where $u(a_i)$ and $u(a_j)$ are the uncertainties associated with the activities of the radionuclides i and j, $r(a_i, a_j)$ is the correlation between their activities and AL_i and AL_j are the activity limits of the radionuclides i and j. A simple example follows.

Example 10 *A package of activated copper is characterized via direct γ-spectrometry, scaling factors and the mean activity method. Co-60 is the only ETM radionuclide found and its activity is 1 Bq/g with a relative uncertainty u_c(a(Co-60)) = 5% (k=1). The activity of Co-60 is correlated with that of Ni-63 and their correlation is r(a(Co-60), a(Ni-63)) = 0.8. Applying the scaling factor method we found SF=2. The combined relative uncertainty of the activity of Ni-63 is u_c(a(Ni-63)) = 40% (k=1). The calculation showed also that H-3 is to be expected but no correlation was found with the KN (r(a(Co-60), a(H-3))) $\simeq 0$. The activities of the DTM radionuclides are also uncorrelated. Applying the mean activity method we found an average H-3 content of 5 Bq/g. The relative standard error of the mean is $u_c(\overline{a}$(H-3)) = 10% (k=1). The activities of ITM radionuclides were calculated but their level is systematically below the Declaration Threshold and are not used to calculate the IRAS. For this package the IRAS is:*

$$IRAS = \frac{a(Co\text{-}60)}{AL(Co\text{-}60)} + \frac{a(Ni\text{-}63)}{AL(Ni\text{-}63)} + \frac{a(H\text{-}3)}{AL(H\text{-}3)} = \frac{1}{10} + \frac{2}{1000} + \frac{5}{1000} = 0.107.$$

If we calculate the contributions of each radionuclide to the IRAS we can see that Co-60 contributes for about 93.46%, Ni-63 for about 1.87% and H-3 for about 4.67%. The magnitudes found here are very common for VLLW characterized at CERN.

Applying Eq. 4.16 to our example (we made the hypothesis the the bias terms are set to 0) we can quantify the uncertainty of the IRAS as follows:

$$u_c^2(IRAS) = \frac{u_c^2(a(Co\text{-}60))}{AL^2(Co\text{-}60)} + \frac{u_c^2(a(Ni\text{-}63))}{AL^2(Ni\text{-}63)} + \frac{u_c^2(a(H\text{-}3))}{AL^2(H\text{-}3)} + \\ + \frac{2 \cdot r(a(Co\text{-}60), a(Ni\text{-}63)) \cdot u(a(Co\text{-}60)) \cdot u(a(Ni\text{-}63))}{AL(Co\text{-}60) \cdot AL(Ni\text{-}63)} = \\ = \frac{0.05^2}{10^2} + \frac{0.4^2}{1000^2} + \frac{0.5^2}{1000^2} + \frac{2 \cdot 0.8 \cdot 0.05 \cdot 0.4}{10 \cdot 1000} = \\ = 0.00002861.$$

We can finally evaluate the expanded uncertainty $U(IRAS)$ (k=2):

$$U(IRAS) = 2 \times u_c(IRAS) = 2 \times 0.00535 = 0.0107.$$

The absolute uncertainty corresponds to an expanded relative uncertainty 0.0107/0.107 = 10%. We can finally evaluate the uncertainty budget calculating the contribution of each uncertainty term to the total uncertainty. We obtain that the uncertainty associated with the activity of Co-60 contributes for 87.38%, that of Ni-63 for 0.56%, that of H-3 for 0.87% and, finally, the mixed term for 11.18%.

Once again, these contributions are commonly found in the radiological characterization of legacy waste produced at CERN.

□

The IRAS of a batch of radioactive waste is calculated using Eq. 1.13. The uncertainty of this quantity can be quantified using the uncertainty propagation formula for uncorrelated quantities. This hypothesis can be made because the $IRAS$ and the weight of a waste package are usually not correlated. In particular we obtain the following:

$$u_c^2(\langle IRAS \rangle) \simeq \sum_{k=1}^{n} \left(\frac{M_k}{M_{tot}} \right)^2 \cdot u^2(IRAS_k) + \sum_{k=1}^{n} \left(\frac{IRAS_k}{M_{tot}} \right)^2 \cdot u^2(M_k) \tag{4.17}$$

where M_{tot} is the total weight of the batch of waste and $u^2(IRAS_k)$ and $u^2(M_k)$ are the squared uncertainties of the IRAS and of the weight of the kth package.

If we divide left and right sides of Eq. 4.17 by $\langle IRAS \rangle^2$ we can formulate the uncertainty as follows:

$$\left(\frac{u(\langle IRAS \rangle)}{\langle IRAS \rangle}\right)^2 \simeq \sum_{k=1}^{n} \left(\frac{u(IRAS_k)}{IRAS_k}\right)^2 + \sum_{k=1}^{n} \left(\frac{u(M_k)}{M_k}\right)^2. \tag{4.18}$$

If the uncertainties of the $IRAS$ are similar ($u(IRAS_k) = u(IRAS)$) and so are the uncertainties of the weight of a waste package ($u(M_k) = u(M)$), Eq. 4.17 can be simplified. Using the definition of variance we can write:

$$\sum_{k=1}^{n} M_k^2 = n\overline{M}^2 + nVar(M) \simeq n\overline{M}^2 + ns^2(M) \tag{4.19}$$

and

$$\sum_{k=1}^{n} IRAS_k^2 = n\overline{IRAS}^2 + nVar(IRAS) \simeq n\overline{IRAS}^2 + ns^2(IRAS) \tag{4.20}$$

where s^2 represents the square of the experimental standard deviations of M and $IRAS$, and $\overline{M}$ and $\overline{IRAS}$ are the average weight and $IRAS$.

Injecting last two equations into Eq. 4.17 and rearranging the terms, we can calculate the uncertainty of the IRAS of a batch of waste as follows:

$$\begin{aligned} u_c^2(\langle IRAS \rangle) \simeq & \frac{1}{M_{tot}^2}\left[u^2(IRAS)(n\overline{M}^2 + ns^2(M))\right] + \\ & \frac{1}{M_{tot}^2}\left[u^2(M)(n\overline{IRAS}^2 + ns^2(IRAS))\right]. \end{aligned} \tag{4.21}$$

Finally, recalling that $M_{tot}^2 = n\overline{M}^2$, we can write:

$$\begin{aligned} u_c^2(\langle IRAS \rangle) \simeq & \frac{u^2(IRAS)}{n}\left(1 + \frac{s^2(M)}{\overline{M}^2}\right) + \\ & \frac{u^2(M)}{\overline{M}^2}\left(\frac{\overline{IRAS}^2}{n} + \frac{s^2(IRAS)}{n}\right). \end{aligned} \tag{4.22}$$

Summary

In this chapter we presented the strategies adopted to evaluate the uncertainties and the bias associated with the characterization process of VLLW at CERN.

We discussed dedicated methodologies for ETM, DTM and ITM radionuclides.

The principal component of the uncertainty of ETM's activity derives from the efficiency calibration function. Activity and weight distributions within a waste package are often unknown and cannot be easily estimated. A number of mitigating actions are taken into account to limits the effect of the bias term.

The uncertainty of the DTM's activity is mainly due to the uncertainty of either the scaling factor or the mean activity of the collected samples. The sampling stage has a considerable impact on the quality of the scaling factor and must be carefully undertaken.

Finally, the uncertainty of ITM's activity is related to the activation studies performed to evaluate the radionuclide inventory and the correlation factors. The estimation of both the random uncertainty and the bias term is difficult. We implemented a Monte Carlo method to quantify this component of the IRAS uncertainty.

Combining the uncertainties associated with the activities of the three families of radionuclides we finally established a formula to quantify the uncertainties of the IRAS and of the weighted IRAS of a waste population.

Chapter 5

Case study, results and discussion

Prediction is very difficult, especially if it's about the future!

Attributed to Niels Bohr
As quoted by A.K. Ellis.

Introduction

The present chapter is divided into three major sections. In the first section we apply the characterization process to a simple, low-heterogeneous waste population of shredded copper from activated cables. The study of the shredded copper population is used to show in detail how the various techniques of characterization can be applied to a real waste population. We will also describe in detail each step of the characterization process according to the flux diagram of Fig. 1.10, Ch. 1.

The second part is dedicated to the analysis of the results obtained during the major disposal campaigns performed in 2015-2017 at CERN, in the context of the so-called SHERPA project [Algoet et al. 2015]. We illustrate here the industrialization of the characterization process. At the date of writing (June 2017) the characterization method has been used to characterize more than 690 tons of metallic legacy waste.

The third section discusses various aspects of the characterization process including cost and time of sampling and measurements, limitations of the methodology proposed, ongoing improvements and future work.

5.1 Radiological characterization of copper from shredded cables

Activated cables represent about 20% of the waste presently stored, or soon to be stored, at the CERN's temporary storage [La Torre and Magistris 2016]. For this reason various studies were performed at CERN in the past years to radiologically characterize this family of waste [Zaffora et al. 2016][La Torre and Magistris 2016][Magistris et al. 2017a][Vincke and Theis 2015].

In this section we focus on a sub-population of activated cables made of copper. After a general description of the cable population we discuss the evaluation of the radionuclide inventory. We show the results of measurements to quantify the activity of ETM radionuclides and we discuss the application of scaling and correlation factor methods to quantify the activity of DTM and ITM radionuclides. We end the section evaluating the IRAS of the waste population and its uncertainty, comparing those values with the acceptance limits for very-low-level waste.

5.1.1 The waste population

General considerations

Cables at CERN are used for a wide range of applications including signal transmission, power distribution, data dispatching and accelerators control. According to their specific use cables can be of different materials and sizes. Their composition can change also depending on the year of production and the specific application they were designed for.

The major elements of cables are the insulation, which separates electrically and physically conductors, the jacket which physically protects the internal components of a cable and provides flame retardancy, and the central metallic core which is dedicated to electric and signal transmission.

Common materials used for insulation and jackets are the Polyvinyl Chloride (PVC) and Polyethylene (PE). To ensure the mechanical characteristics of PVC (flexibility, transparency, durability and longevity) but also to increase its resistance to degradation, a number of substances can be added. We can cite two major categories of additive: stabilisers

and plasticisers [Chabrol and Girois 2013] [Verrier 1992]. Common stabilizers of PVC are lead compounds, organotin compounds, mixed metals and organic based stabilizers [Chabrol and Girois 2013]. Common plasticisers are phthalate ester, epoxides, polyester plasticizers and phosphates [Verrier 1992].

Metallic cores of cables are commonly made of copper and aluminium but other metals can be found, according to the specific application.

When combining the various types of insulations, jackets and metallic cores, theoretical limitless cable types can be produced. The constitutive elements of cables are however limited and can be grouped according to their similarities in terms of material and/or type.

At CERN very-low-level radioactivity cables are presently classified into 5 major families according to the amount of plastic and the type of metallic core they are made of [La Torre and Magistris 2016]. This classification is not meant to be representative of all possible cables activated at CERN but it allows us to sort the waste according to easily recognizable criteria (i.e. type of core and insulation/jacket).

The 5 families consist of power cables with copper or aluminium central wires, copper signal, copper coaxial and mixed copper/aluminium coaxial cables. Figs. 5.1-5.3 show some examples of cables drawn from the 5 families. Tab. 5.1 gives the average density of the cables (core plus insulation/jacket) and the average amount of metal for each of the cable families previously described. The uncertainty associated is calculated as the standard error of the mean values obtained from multiple samples [La Torre and Magistris 2016].

Table 5.1: Average density and average amount of metal of the 5 families of VLLW cables identified at CERN. The uncertainty is calculated as the standard error of the mean (k=1) [La Torre and Magistris 2016].

Family	**Average density (g/cm^3)**	**Average metal mass (wt-%)**
Al power	1.9 ± 0.1	53.5 ± 6.9
Cu power	2.8 ± 0.6	63.8 ± 5.4
Cu signal	1.6 ± 0.1	47.3 ± 1.5
Cu coaxial	1.7 ± 0.1	42.4 ± 4.2
Cu/Al coaxial	1.3 ± 0.2	46.6 ± 14

Figure 5.1: Sample CR-049963 of aluminium power cable ($\phi = 1.8$cm). Photo credit: F.P. La Torre, CERN.

Figure 5.2: Sample CR-049985 of copper signal cable ($\phi = 2$cm). Photo credit: F.P. La Torre, CERN.

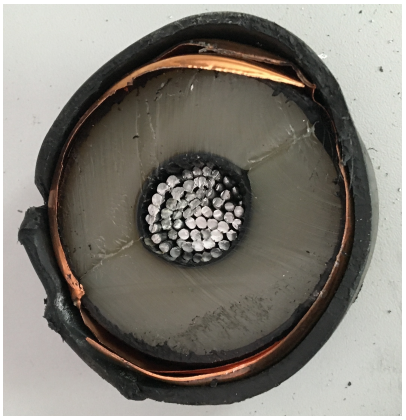


Figure 5.3: Sample CR-049721 of copper/aluminium coaxial cable ($\phi = 6.5$cm). Photo credit: F.P. La Torre, CERN.

Description of the population

To test and validate the characterization process described in the previous chapters, we selected a waste population made of legacy copper cables which were separated from the insulation/jacket and shredded. Very limited information is available about this population in terms of radiological history. Preliminary qualitative measurements performed on chosen samples indicate that the concentration of the element Cu is above 99%. We will refer to the material type as generic cathodic copper.

The metallic core and the insulating plastic layers were separated and shredded to diminish the heterogeneity of both the elemental composition and the activity. A trimmer-shredder, with a mixing capability, was used for this purpose. Fig. 5.4 shows an example of shredded copper from the waste population considered. The granularity of the shreds can change depending on the original copper type, but the grain size is consistently below 0.2 mm.

Figure 5.4: Example of shredded copper.

The shredded copper was stored in 87 drums (60 L capacity). The average weight of the drums is 97 kg and the waste population weight is 8522 kg. Power, signal and coaxial cables are part of the waste population. We give the results of a preliminary γ-screening performed with a total-γ counter in Tab. 5.2.

Each drum was measured and the specific activity is expressed as Co-60 equivalent, that means that each detected photon is considered as being emitted by a Co-60 source.

Table 5.2: Summary statistics of the preliminary γ-screening performed on the copper waste population. The reference date for the Co-60 equivalent specific activity is 01/07/2015.

Statistics	**Value (Bq/g)**
Mean Co-60 activity	0.041
Standard error of the mean at k=1	0.002
Median Co-60 activity	0.038
I.Q. range	0.025
Min Co-60 activity	0.005
Max Co-60 activity	0.138

The reference date for the activity is 01/07/2015. The calibration of the system assumes that the mass and activity distributions within the drums are uniform. Various geometric calibrations are however used to take into account the filling levels of the sampling units.

5.1.2 Calculation of the radionuclide inventory

The waste population is made of a mix of shredded copper from cables of unknown radiological history. Preliminary qualitative measurements performed on a limited sample indicate that the copper content is above 99% but trace elements cannot be quantified with the instruments commonly used at CERN[1].

In such an uncertain context, we decided to include within the elemental composition used for the calculation of the radionuclide inventory, a large list of potential impurities and contaminants as described in Ch. 2. We finally performed a sensitivity study to evaluate the impact of each input parameter on the final results using the statistical learning tools introduced in Ch. 3. The list of impurities consists of the elements found in 119 samples from ores collected in representative mining sites around the world by the European Copper Institute [Delbeke and Rodriguez 2014]. A detailed description of the distribution of each element can be found in Ch. 2, Tabs. 2.4-2.6.

The radiological history of the activated cables being unknown, we decided to consider as equally probable the machines of origin, the positions within the tunnel accelerator and, the irradiation and decay times from 0 up to 40 years. Of course some combinations of input parameters are improbable but, we would like to recall here that these preliminary

[1] An instrument commonly employed at CERN for fast elemental analyses on waste is the XRF analyzer commercialized by ThermoFisher Scientific (`https://www.thermofisher.com/ch/en/home.html`, February 2017). As an example, the minimum detectable concentration of the element Co in metals is above 50 ppm.

calculations are needed to generate an envelope case of radionuclides that can be produced by activation at common CERN's radiological environments.

For the purpose of the present study we extracted 10000 realizations from the random vector **S**. We showed in Ch. 2, Sec. 2.5 that this number of realizations is a good compromise between computer simulation time and accuracy.

We compared then the simulated activity of each radionuclide produced in each scenario with its Declaration Threshold and, if the activity was above the limit, the radionuclide was included within the preliminary inventory (see Ch. 3, Sec. 3.3, Eq. 3.23).

The preliminary radionuclide inventory obtained for the waste population of shredded copper is given in Ch. 2, Tab. 2.7. The table is reproduced here for clarity.

Radionuclide	**ETM**	**DTM**	**ITM**	**C_{IRAS} (in %)**
Co-60	✓			93.91
Ni-63		✓		3.17
H-3		✓		2.27
Ag-108m	✓			0.35
Ti-44	✓			0.20
Fe-55			✓	0.05
Total				99.95

A number of other radionuclides is predicted by the calculations. Their contribution to the IRAS and their activity are however so low that they can be neglected for the analyses that follow. It should nevertheless be noticed that the γ-emitters are measured in every single waste package. Any discrepancy between calculations and measurements can therefore be identified easily.

5.1.3 Quantification of ETM's activity

According to the preliminary radionuclide inventory, the major ETMs expected in the waste population are Co-60, Ag-108m and Ti-44. A systematic γ-screening is however able to detect the presence of radionuclides which were not expected.

To minimize the impact of the potential heterogeneities within each sampling unit, we collected three samples at three different levels of each drum, and mixed them generating a composite sample (see Ch. 3, Sec. 3.2.3 for details). Samples of copper shreds collected for

the characterization of the waste population are showed in Fig. 5.5.

Figure 5.5: Collected samples of copper shreds.

The sample collected is a census. Census data is rarely collected for waste characterization because of the cost and time that such a sampling method would require.

After mixing, each composite sample was split into three sub-samples: one sub-sample to be measured via γ-ray spectrometry for ETM's activity quantification, one sub-sample to quantify the activity of the DTMs Ni-63 and H-3 and the ITM Fe-55 and, finally, one sub-sample to be stored as a reference. This operation was repeated for the entire waste population.

The weight of the samples spans between 25 and 80 grams. The counting time chosen to measure the 87 samples via γ-spectrometry is 10000 seconds. This counting time is a good compromise between the achievable MDA ($< 2 \cdot 10^{-2}$ Bq/g for Co-60) and the uncertainties (the median relative standard deviation for the activity of Co-60 is 8%). The two most radioactive samples were measured for a longer counting time (48 hours) in order to measure specific activities close to the Declaration Threshold of Ag-108m (DT=$2.5 \cdot 10^{-4}$ Bq/g).

A dedicated efficiency calibration was performed for each sample according to the apparent density and filling level of the sample. The reference date to report the specific activity is the first of July 2015. 81 samples out of the 87 collected showed a specific

activity of Co-60 above the MDA of the spectrometer used. Summary statistics are showed in Tab. 5.3.

Table 5.3: Summary statistics of the measurements performed on 87 composite samples of shredded copper. The reference date for the specific activity is 01/07/2015. Last line indicates the Declaration Thresholds of the various nuclides.

Radionuclide	**H-3**	**Co-60**	**Ni-63**
Samples above MDA	80	81	64
MDA (Bq/g)	0.02-0.06	0.005-0.02	0.08-0.3
Min activity (Bq/g)	0.03	0.009	0.16
1^{st} quartile (Bq/g)	0.22	0.031	0.43
Mean activity (Bq/g)	0.45	0.046	0.65
Median activity (Bq/g)	0.36	0.042	0.55
3^{rd} quartile (Bq/g)	0.51	0.058	0.71
Max activity (Bq/g)	5.68	0.137	2.31
IQ range (Bq/g)	0.29	0.027	0.28
Standard error of the mean (Bq/g)	0.07	0.003	0.05
Skewness	6.89	1.02	2.34
Kurtosis	56.46	4.47	9.86
Declaration Threshold (Bq/g)	1	0.1	10

The analyses performed on the samples show the presence of some traces of Ti-44 (one sample above the MDA), Co-57 (three samples above MDA) and Mn-54 (two samples above MDA). These radionuclides, together with the KN Co-60, represent the list of ETMs for the waste population under consideration. Their presence was predicted by the calculations but their activity is systematically below the Declaration Threshold. According to reference [ANDRA 2013a] their activity should not be declared or used for IRAS calculation. Co-60 will be used to test the applicability of the scaling factor method and evaluate the activity of DTM radionuclides.

Fig. 5.6 shows the distribution of the specific activity of Co-60 in the 87 drums and its uncertainty. Data below the Minimum Detectable Activity was replaced by the corresponding MDA and its error is stated as MDA/2. From the figure it is easy to see that the level of Co-60 radioactivity is very-low and only very few packages have a Co-60 activity above the DT (in red in figure). The average Co-60 activity is showed in green.

The majority of the relative uncertainties of the activity of Co-60 is below 15% (k=1) with a median value of 8%. It should be outlined that the level of the γ activity of the waste population is often close to background level making the discrimination between

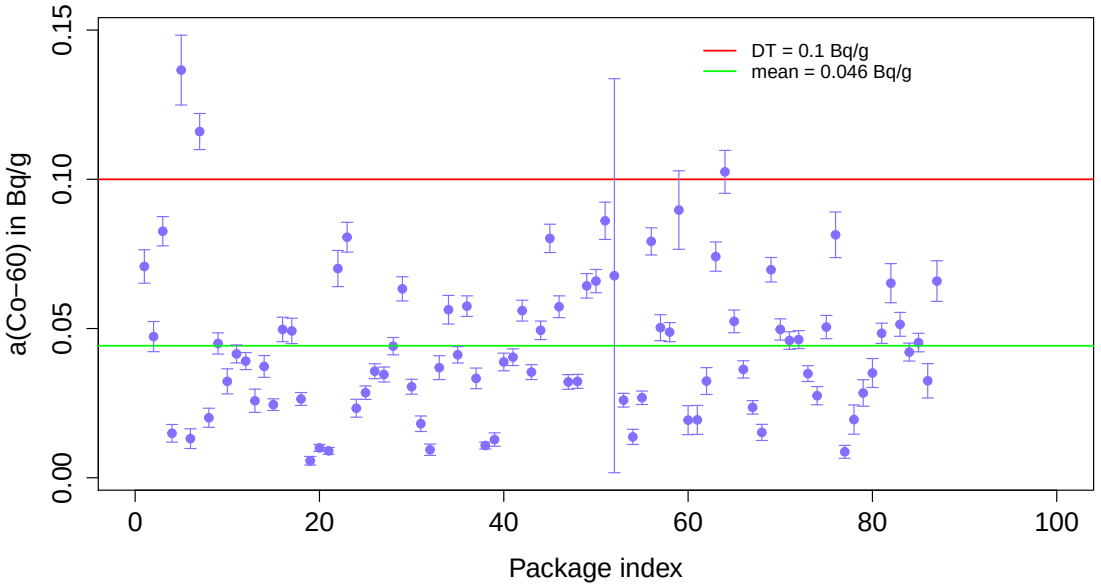


Figure 5.6: Specific activity distribution of Co-60 in the waste population made of 87 drums of shredded copper from cables. The error bars are given for a coverage factor k=1. For the 6 packages with specific activity below the MDA, the specific activity is replaced by the MDA and the uncertainty is calculated as MDA/2. The green line indicates the average Co-60 specific activity for the entire population. In red it is showed the Declaration Threshold of Co-60 (0.1 Bq/g).

the signal and the background difficult. The direct impact of such a low activity can be seen on the relatively high uncertainty showed. For higher activities, the relative standard uncertainty commonly lies below 5% (for k=1).

In one case, the relative standard deviation is 97.5% (sample CR-045063, package index 52). This extreme value is due to a very low γ signal which cannot be properly separated from the background. The specific activity is however quantified above the MDA and its uncertainty is very high.

5.1.4 Quantification of DTM's activity

The list of the major radionuclides predicted via simulations includes two DTMs which contribute above 1% to the IRAS. These radionuclides are Ni-63 and H-3.

In Ch. 3 we discussed the scaling factor method to evaluate the activity of DTM

radionuclides. A general rule for the application of the scaling factor technique implies that a consistent correlation must be tested between the specific activity of the key nuclide, such as Co-60, and the specific activity of DTMs, over a representative sample withdrawn from the population.

Tab. 5.3 shows that 64 and 80 samples for, respectively, Ni-63 and H-3 have a specific activity above the MDA. Using the values above the MDA we found that only the pair Co-60/H-3 has a correlation coefficient above 0.5 (r=0.51). The geometric scaling factor could be used to quantify the activity of H-3 in the waste population. For the pair Co-60/Ni-63 the correlation coefficient is below 0.5 (r=0.33). In this case the geometric scaling factor is not suitable to be used and alternative methods, such as linear models or the mean activity method, can be used.

The experimental geometric scaling factor for H-3 is $\overline{G}_{SF}^{h3} = 8.67$ (see Eq. 3.10). Its standard uncertainty is $u(\overline{G}_{SF}^{h3}) = 0.55$ for a coverage factor k=1 (see Eq.4.12).

We further built linear models to express the relationship between the KN and the DTM radionuclides. The linear models are built using only the values above the MDA. If the intercept coefficient $\hat{\beta}_0$ of Eq. 3.8 is set to 0, we obtain the following relationships: $\hat{a}_{ni63,i} = 11.32 \times a_{co60,i}$ ($R^2 = 0.74$) and $\hat{a}_{h3,i} = 11.42 \times a_{co60,i}$ ($R^2 = 0.5$). The exploratory data analysis of the linear models built is showed in Figs. 5.7-5.8.

Fig. 5.7 shows the exploratory data analysis of the linear model predicting the activity of H-3. We can easily see that the model without intercept is more robust ($R^2 = 0.5$). From the residual plot, histogram and normal qq-plot it is also possible to see the existence of an outlier. In absence of physical evidence or reasons to exclude the outlier, this value should be considered for the statistical analysis. The data is considerably homogeneous, the residuals are homoscedastic and the normal approximation seems reasonable.

As a consequence of the previous analysis the linear model (with $\hat{\beta}_0 = 0$) can also be used to predict the activity of H-3 from the measured activity of Co-60.

The analysis of the exploratory data plots of the linear models of the pair Ni-63 vs Co-60 (see Fig. 5.8) indicates that the linear model obtained from regression (and $\hat{\beta}_0 = 0$) is appropriate also for Ni-63.

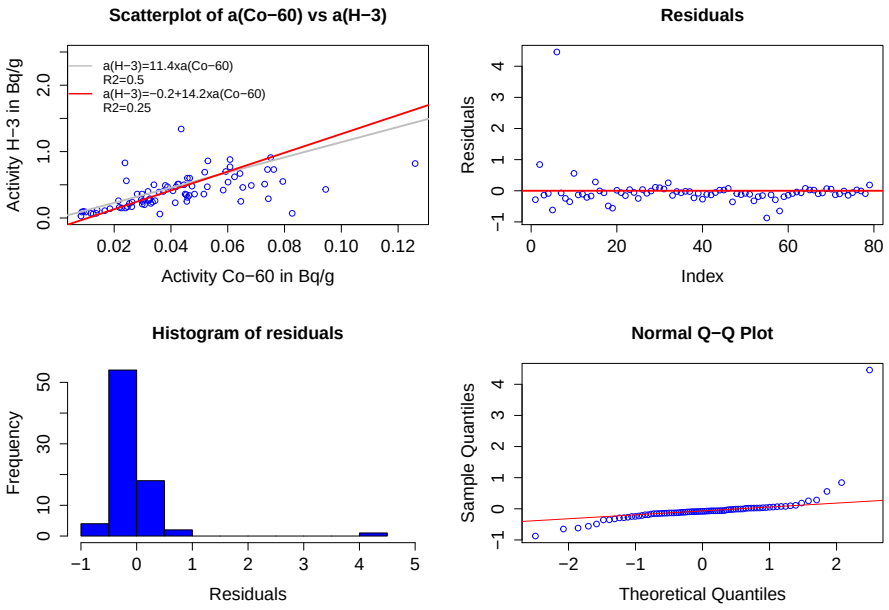


Figure 5.7: Study of the linear model of H-3 vs Co-60. The first plot shows the scatterplot of the activities together with the regression lines with $\hat{\beta}_0 = 0$ (gray) and $\hat{\beta}_0 \neq 0$ (red). The plots 2-to-4 present the dispersion, the histogram and the normal qq-plot of the residuals.

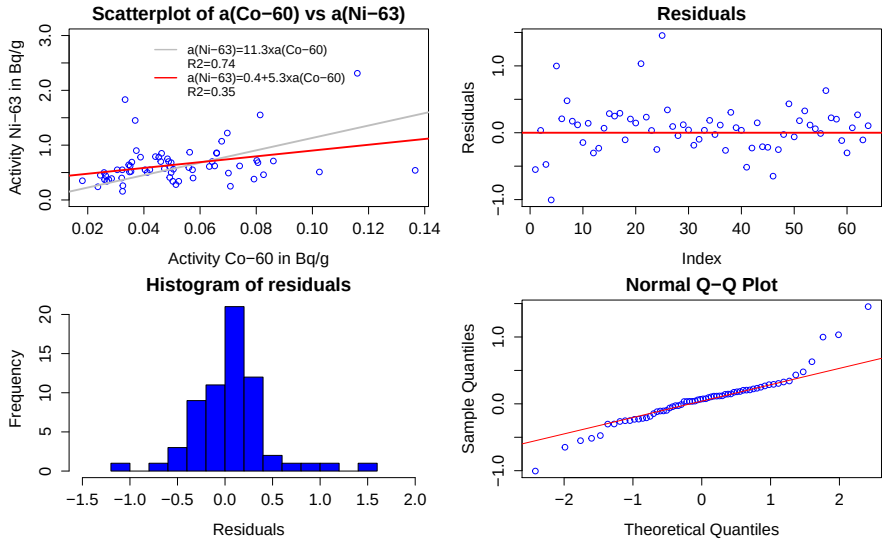


Figure 5.8: Study of the linear model of Ni-63 vs Co-60. The first plot shows the scatterplot of the activities together with the regression lines with $\hat{\beta}_0 = 0$ (gray) and $\hat{\beta}_0 \neq 0$ (red). The plots 2-to-4 present the dispersion, the histogram and the normal qq-plot of the residuals.

Using the measured activities of the key nuclide Co-60, the experimental geometric scaling factor for H-3 and the linear model for Ni-63, we calculated the specific activities of H-3 and Ni-63 in the waste population. The results are presented in Figs. 5.9-5.10.

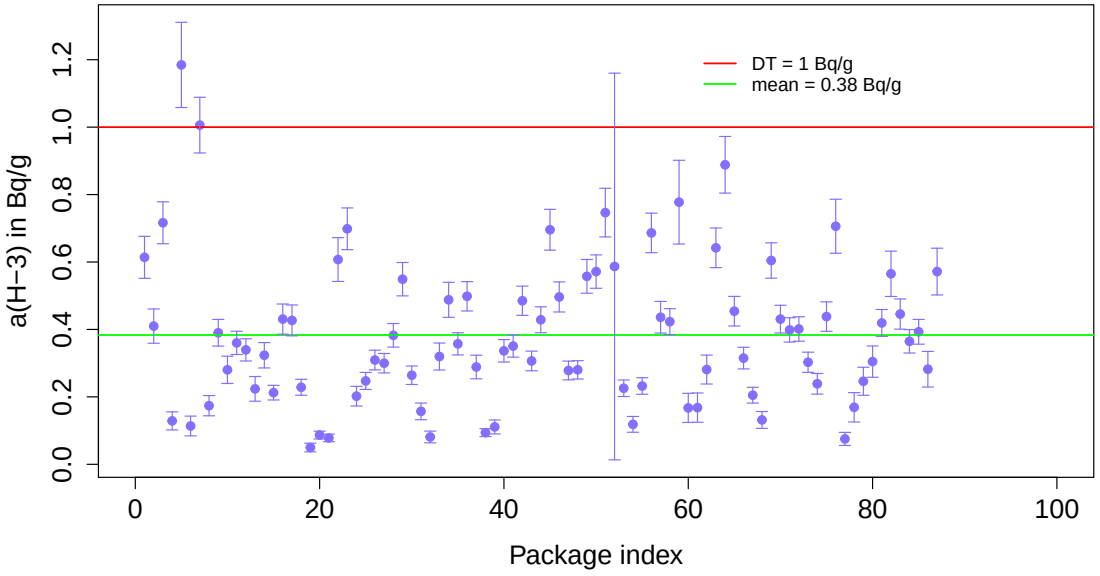


Figure 5.9: Activity distribution of H-3 in the waste population made of 87 drums of shredded copper from cables. The error bars are given for a coverage factor k=1. In green is showed the average activity (0.38 Bq/g) and in red the Declaration Threshold (1 Bq/g).

From Fig. 5.9 we can see again that only a very limited number of packages have an activity level above the Declaration Threshold of H-3 (DT = 1 Bq/g). Finally, Fig. 5.10 shows clearly that the specific activities of Ni-63 in the waste population are one order of magnitude below the Declaration Threshold.

5.1.5 Quantification of ITM's activity

Although in routine characterization ITM radionuclides are not sampled at CERN, in this study we measured also the specific activity of the major ITM. The first ITM appearing in the calculated radionuclide inventory is Fe-55, which contributes to the IRAS up to 0.05% (see Tab. 2.7). The scope of this complementary analysis was to check the reliability of the algorithm to discriminate DTMs from ITM (see Eq. 2.2). Only three samples, out of the

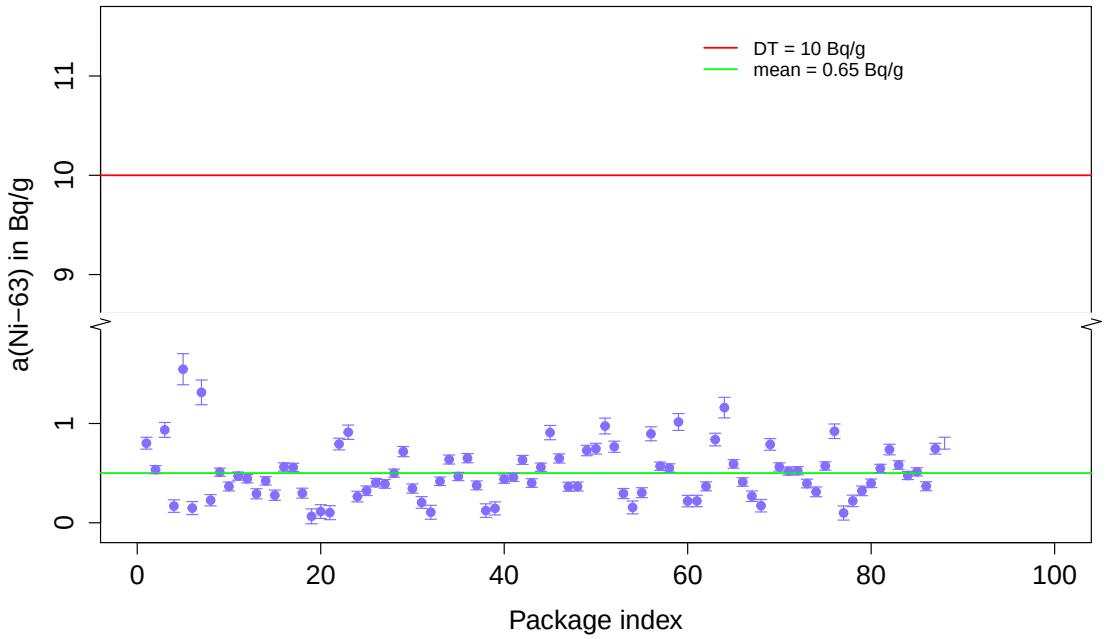


Figure 5.10: Activity distribution of Ni-63 in the waste population made of 87 drums of shredded copper from cables. The error bars are given for a coverage factor k=1. In green is showed the average activity (0.65 Bq/g) and in red the Declaration Threshold (10 Bq/g).

87 withdrawn from the waste population, have an activity above the MDA (0.15-1.1 Bq/g). The corrected activities of Fe-55 for the three samples at the reference date (01/07/2015) are 0.38 Bq/g (CR-037369), 0.78 Bq/g (CR-037445) and 0.89 Bq/g (CR-037448). These specific activities are significantly below the Declaration Threshold (DT(Fe-55)=10 Bq/g).

The agreement between the census data of the case study and the routine measurements performed at CERN indicates that the quantification of ITM radionuclides via radiochemical analyses is not justified from both a technical and a financial point of view.

Since a sample is available but not enough data is above the MDA, we can also apply the mean activity method to evaluate the activity of Fe-55 in the waste population. Using the mean activity method we found $\overline{a}_{fe55} = 0.38 Bq/g$. The standard error of the mean is 0.17 Bq/g for a coverage factor k=1. Using the results from the calculations performed (10000 realizations of the vector scenario) we find a geometric correlation factor $\overline{G}_{CF}^{fe55} = 0.054$. Independently from the method chosen to estimate the specific activity, the amount of Fe-55 is systematically below the Declaration Threshold from one (mean activity) to three

(geometric correlation factor) orders of magnitude.

The same reasoning is followed to establish that no other ITM radionuclide should be considered to calculate the IRAS of the waste population.

5.1.6 Calculation of IRAS

Using Eq. 1.11 and the activities presented in the previous sections we can finally calculate the IRAS of the waste packages. The results are presented in Fig. 5.11. We found that the average IRAS of the 87 packages is 0.005, well below the acceptance limit 10 for VLLW waste in France.

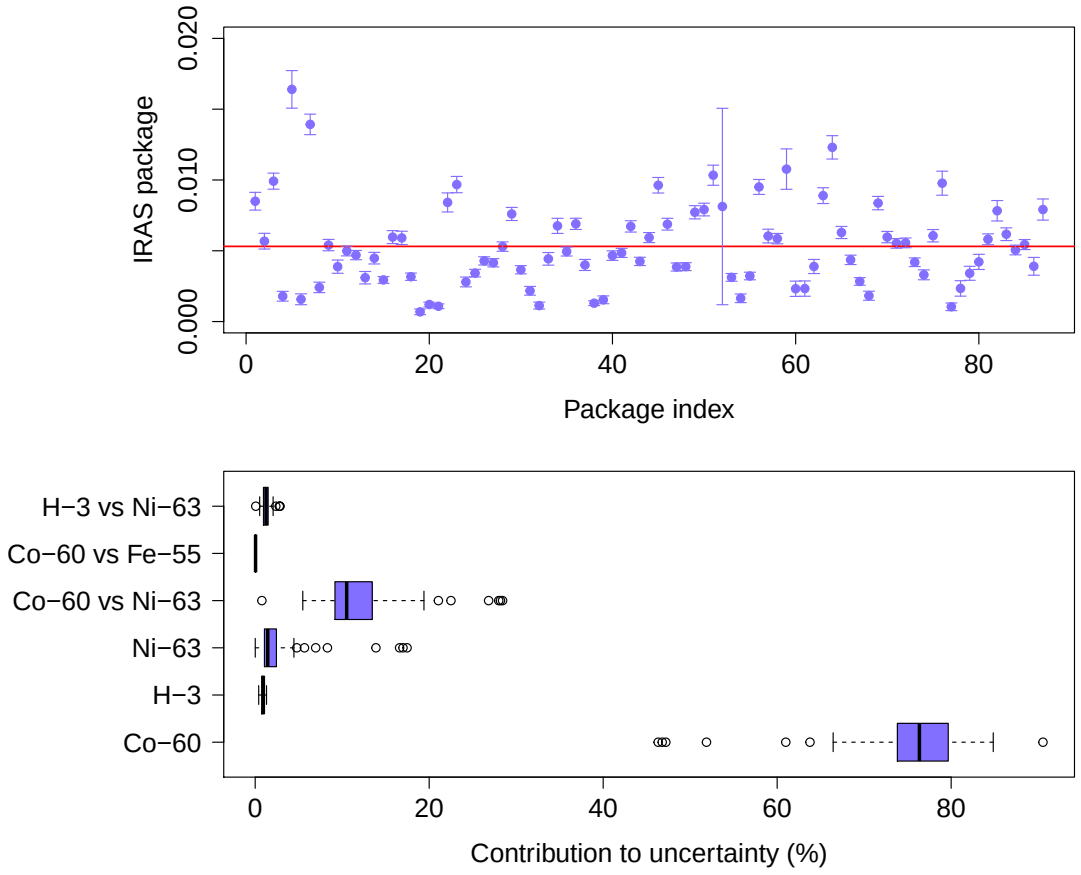


Figure 5.11: The top plot shows the IRAS and its standard deviation (k=1) for the 87 waste packages of very-low-level radioactive copper activated at CERN. The bottom plot illustrates the distributions of the contribution of the uncertainty of each radionuclide's activity to the total IRAS uncertainty. Radionuclides whose activity's uncertainty contributes less than 0.1% are omitted for clarity.

The general formula to express the combined standard uncertainty of the IRAS of a waste package is given in Eq. 4.16. For completeness, Tab. 5.4 presents selected statistics of the relative standard deviations of the activities of the major radionuclides in the waste population. The coverage factor is k=1.

Table 5.4: Summary statistics of the relative standard deviation u_{rel} (in %, k=1) associated with the activities of the major radionuclides in the copper waste population.

Statistics	$\mathbf{u_{rel}(a_{co60})}$	$\mathbf{u_{rel}(a_{h3})}$	$\mathbf{u_{rel}(a_{ni63})}$	$\mathbf{u_{rel}(a_{fe55})}$
Minimum	5.2	8.2	7.1	4.8
1^{st} quartile	7.0	9.5	7.5	7.1
Median	8.1	10.4	9.5	8.3
Mean	11.5	13.4	17.1	11.9
SE of the mean	1.2	1.1	2.0	1.2
3^{rd} quartile	12.8	14.4	16.6	12.9
Maximum	97.5	97.7	115.4	100.0
Skewness	6.0	6.4	2.9	5.5
Kurtosis	44.3	48.5	9.5	38.1

The top plot of Fig. 5.11 shows the IRAS distribution in the waste population. The average relative standard uncertainty of the IRAS is 11% (k=1). The bottom plot of Fig. 5.11 shows that the uncertainty associated with the activity of the key nuclide contributes the most to the standard uncertainty of the IRAS. We found that the average contribution of the uncertainty of Co-60 is above 75%. If we add also the contributions of the combined Co-60 terms, the contribution to the total uncertainty of the IRAS adds up to ∼90%. For the waste population considered in this study the average contribution of the uncertainties associated with the activities of DTM and ITM radionuclides alone are below 2.5%. The terms contributing less than 0.1% to the total uncertainty of the IRAS are omitted from the analysis.

The reduction of the uncertainty of the IRAS is therefore related to the reduction of the uncertainties associated with the activity of the key nuclide. This consideration however cannot be generalized to other waste populations with a different radiological history. In other cases the activity's uncertainty associated with DTM and ITM radionuclides can play an important role.

We can finally calculate the IRAS of the batch using Eq. 1.13 and quantify its uncertainty according to Eq 4.17. We found that $\langle IRAS \rangle = 2.5 \times 10^{-4}$ and that its relative standard uncertainty (for k=1) is ∼1.5%. The IRAS of the batch is well below the acceptance limit.

Even if the activities so calculated generate IRAS which are well below the acceptance limits, clearance from regulatory control is not allowed in France where all materials originating from nuclear zones must be considered as radioactive.

The activity levels found in this case study are common at CERN. The choice to dispose of materials as radioactive waste or clear them from regulatory control is regulated by the Tripartite Agreement [Acc 2011] and it is made on the basis of a fair share policy between the Host States, together with concern for technical difficulties and time/cost constraints.

In this case study we had the opportunity to use many data which, as a general rule, are not available for operational waste characterization. The method presented in this case study is today partly industrialized for the characterization of metallic waste. The summary of the disposal campaigns performed at CERN in 2015-2017 is the topic of next section.

5.2 The SHERPA project

This section presents the summary of the results obtained in the context of the so-called SHERPA project, launched at CERN end of 2014. The project represents the first industrial-scale application of the characterization process introduced in this thesis.

We first describe the aim of the SHERPA project and outline the characteristics of the waste families treated. We then present the calculation of the preliminary radionuclide inventory and illustrate the quantification of the activities of ETM, DTM and ITM radionuclides. We conclude the section with the calculation of the IRAS and the summary results after two years of operations.

5.2.1 The waste population

General considerations

The goal of the SHERPA project is to set-up an effective and sustainable compressing and shearing process for substantial quantities of metallic radioactive waste at the CERN radioactive waste treatment centre [Algoet et al. 2015].

To this end, the project has:

1. identified in the radioactive waste storage 1000 tons of metallic VLLW for processing;

2. obtained the approval from ANDRA for the disposal of the 1000 tons of metallic waste over 5 years;

3. documented the procedures, quality assurance and quality control for the characterization of the radioactive waste packages;

4. defined the operational procedures to set-up a safe and effective compressing and shearing process; and

5. processed and disposed of a pilot campaign of 200 tons of metallic VLLW; at the date of writing (June 2017) above 690 tons of VLLW has been characterized and partly disposed of in the disposal facility (above 80% is already disposed of at the final repository).

The treatment of the metallic waste relies on a 600 tons industrial press-shears (see Fig. 5.12 [Bruno et al. 2017]), capable of compacting and cutting 2 m-wide 5.9 m-long metallic radioactive items to a volume and size suitable to be packaged in standard ANDRA containers of 1.3 m^3 and 2.7 m^3 capacity and 200 litres drums with an average density 1 ton per m^3.

Objects commonly treated are pipes, vacuum chambers and structural elements of irregular shapes (see Ch. 2, Fig. 2.2 for an example). Massive items are also included on the disposal program even if they are not treated using the press-shears and are conditioned on the waste package without further treatment. An example of massive objects treated within the frame of the SHERPA project is showed in Fig. 5.13.

The most common metals encountered when characterizing legacy waste produced at CERN are steel, aluminium and copper but limited amounts of other metals can be present at a very low percentage in mass. The SHERPA project deals with metals activated at both hadron (protons and heavy ions) and electron-positron machines. Ad hoc characterization studies must be performed for metals activated at electron-positron machines, such as the Large Electron-Positron collider (LEP). In this thesis we mainly focused on activation

Figure 5.12: Processing of a radioactive waste item by the press-shears [Bruno et al. 2017].

occurring at hadron machines. We will however summarize also the characterization of legacy activated metals originating from the LEP.

Description of the population

After treatment, the waste population is mainly packaged within 1.3 m^3 ANDRA containers as showed in Fig 3.4 of Ch. 3. Starting from 2016 also 2.7 m^3 packages are used for conditioning radioactive waste. Their use was mainly foreseen to contain items of a size not compatible with the dimensions of 1.3 m^3 packages, but are now used also for waste of lower dimensions. The summary of the packages disposed of (>80%) or ready to be disposed of, as of June 2017, is showed in Tab. 5.5.

The origin of the waste treated in the context of the SHERPA project is discussed in Sec. 1.3.3, where the major dismantling campaigns were discussed. We want to add here that, as a general rule, when a new disposal campaign starts, the perimeter of the waste population is clearly identified. This means that, as far as possible, waste originating from one dismantling site is not mixed with waste from another site. The definition of the perimeter of a waste population allows us to identify a so-called campaign. For instance

Figure 5.13: Example of massive objects disposed of as VLLW waste in 2016.

Table 5.5: Summary of the SHERPA disposal campaigns as of June 2017. Above 80% of the packages is already stored at the ANDRA disposal facilities. The remaining waste packages are characterized and ready to be shipped. About 20% of the waste originate from electron machines.

Material	**N. packages**	**Weight (tons)**
Aluminium	68	81.6
Copper	17	38.7
Steel	262	572.4
Total	347	692.7

the summary given in the previous table is calculated from 8 different disposal campaigns performed between 2015 and 2017.

The size of a campaign can be variable depending on the amount of waste with similar characteristics. As an example, the smallest campaign includes 9 final packages for a total weight of ∼16 tons (campaign C008) while the biggest campaign includes 105 waste packages for over 226 tons (campaign C004).

For the campaigns of bigger size it is also common that the production of the packages is made over many months. In this case, the disposal is made after a batch of a fixed size is ready for the expedition at the disposal site. The loading of a package for the transport is showed in Fig. 5.14 [Bruno et al. 2017].

Next section summarizes the calculation of the preliminary radionuclide inventory performed for the SHERPA project.

Figure 5.14: Loading of standard packages filled with metallic VLLW ready for transport to the ANDRA disposal facility [Bruno et al. 2017].

5.2.2 Calculation of the radionuclide inventory

Hadron accelerators

The calculation of the radionuclide inventory for the SHERPA project lies on extensive analytical calculations (over 2.35 million scenarios considered), performed considering 43 reference chemical compositions for steel, aluminium and copper. The tools used to perform these numerical experiments are described in Ch. 2 and Ch. 3.

A difference exists however between the extraction of random realizations from the vector scenario, as suggested in this thesis, and the calculation of the radionuclide inventory as applied to the SHERPA project. In Ch. 2, the radionuclide inventory is calculated from random extractions of elemental compositions, energies, positions, irradiation and decay times. Before 2017, when the new methodology was designed, we calculated the radionuclide inventory considering each possible combination of the input parameters, hence the 2.35 million scenarios.

In practice, we considered all the elemental compositions of metals from the CERN material catalogue [Froeschl et al. 2012] listing the most common materials used for

building accelerator's components and surrounding structures. For each material type, we fixed a dense grid of the remaining input parameters (with irradiation and decay times spanning from 1 year up to 40 years, spaced by 1 year) and performed the calculations for each node of the grid. Finally, a selected extraction from the database is used to estimate the radionuclide inventory of a specific batch of waste. For instance, the decay time commonly considered for characterization studies spans between 3 up to 30 years.

A summary of material families, density ranges, number of grades and chemical elements (major and trace elements) considered is presented in Tab. 5.6.

Table 5.6: Summary data of the major material grades considered for the calculation of the radionuclide inventory for the SHERPA project.

Material	**Density (g/cm^3)**	**Grades**	**Chemical elements**
Aluminium	2.65-2.73	7	Al, Cr, Cu, Fe, Mg, Mn, Na Ni, Si, Ti, V, Zn, Zr
Copper	7.6-8.94	20	Al, B, Be, Bi, C, Cd, Co, Cr Cu, Fe, Hg, Mn, Ni, O, P, Pb S, Si, Sn, Zn, Zr
Steel	7-8.75	16	Al, B, Be, C, Ca, Co, Cr Cu, Fe, Mg, Mn, Mo, N, Nb Ni, P, Pb, S, Si, Sn, Ti, V, Zn

For each scenario considered and each produced radionuclide, we applied Eq. 3.23 to test the relevance of the given radionuclide. Radionuclide inventories estimated for selected families of metals are given in Tab. 5.7. An operational transposition of this study and a short description of the production mechanisms can be found in reference [Magistris et al. 2015].

Table 5.7: Predicted radionuclide inventories of VLLW metals activated at CERN. Above 2.35 million irradiation scenarios were considered.

Metal family	**ETMs**	**DTMs/ITMs**
Aluminium	Na-22, Al-26, Ti-44, Mn-54 Co-57, Co-60, Zn-65	H-3, C-14, Cl-36, Ar-39 Fe-55, Ni-63
Copper	Na-22, Ti-44, Mn-54, Co-57 Co-60, Zn-65, Mo-93, Rh-101 Ag-108m, Ag-110m, Sn-121m, Sb-125 Gd-148, Hg-194, Bi-207	H-3, C-14, Cl-36, Ar-39 Ca-41, Fe-55, Ni-63
Steel	Na-22, Ti-44, Mn-54, Co-57 Co-60, Zn-65, Mo-93, Nb-93m Nb-94, Tc-99	H-3, Be-10, C-14, Cl-36 Ar-39, Ca-41, Fe-55 Ni-63

A few considerations must be made about the results of Tab. 5.7. First of all, the radionuclide inventories are not meant to be representative of all the possible activation scenarios of CERN. The number of scenarios considered can however guarantee a high degree of predictability for potential radionuclides produced when a given family of material is activated.

The inventories were generated using material compositions from the CERN material catalogue [Froeschl et al. 2012]. For this reason random variations of the chemical elements are not taken into account. New radionuclides can be produced if trace elements were present on a specific waste and not included in the list.

At this stage, a differentiation between DTM and ITM radionuclides is not made. In fact, the same radionuclide can both be a DTM or an ITM depending on its contribution to the IRAS. When performing general predictive calculations of the radionuclide inventory we do not therefore distinguish the two families of radionuclides.

When the characterization process of a specific batch of waste starts, detailed analyses are performed and the radionuclide inventory is recalculated. Tab. 5.7 gives however a robust indication of potential radionuclides generated by activation even before that the disposal campaign begins, allowing us to orientate the characterization effort.

Finally, the inventories are dynamic in the sense that new calculations or data can be taken into account whenever available and their robustness can be easily tested via destructive and non-destructive analyses on activated materials.

Electron accelerators

In 2017, a new version of the calculation tool Actiwiz [Theis and Vincke 2012] was released (version 3). The new version allows us to use particle spectra different from the default hadronic spectra described in Ch. 2 (see Sec. 2.3). The new version accepts particle's spectra generated with the Monte Carlo code Fluka [Battistoni et al. 2006][Ferrari et al. 2005].

We tested the new release of the code to produce various radionuclide inventories for the LEP accelerator. The spectra that we used were produced in the context of the CLEAR

project [Duchemin et al. 2017], for the clearance of the supraconducting cavities of the LEP machine. The beam energy is 104.5 GeV and the secondary particles considered are photons, neutrons, protons and positive and negative pions.

For the calculations we considered 4 representative locations nearby the cavities. The irradiation time spans from three up to 11 years (the LEP was commissioned in 1989 and completely dismantled in 2001). The decay time spans from 16 up to 24 years.

The waste population chosen for the test (C008 of SHERPA) consists of ∼16 tons of pipes made of stainless steel (type 304L). Using the chemical composition of the CERN material catalogue [Froeschl et al. 2012] associated with this grade, we produced the radionuclide inventory showed in Tab. 5.8.

Table 5.8: Predicted radionuclide inventories of VLLW pipes made of stainless steel and activated at the LEP accelerator.

Metal family	**ETMs**	**DTMs/ITMs**
Steel	Na-22, Ti-44, Co-60	H-3, C-14, Cl-36, Ar-39 Ca-41, Fe-55, Ni-63

We can see that the list of radionuclides is shorter with respect to the one obtained for proton machines. This is mainly due to the reduced number of scenarios considered. Without going into details, we should also remember that the activation mechanisms can sensibly differ from hadron and electron accelerators and colliders (for instance, spallation reactions can be predominant for hadron machines and photofission can be preponderant for high-energy electron accelerators).

However, only limited differences in terms of nuclide inventory can be seen between waste produced at protons and electron machines. Major ETMs found in waste from hadron accelerators, such as Ti-44 and Co-60, are also the major ETMs found in waste from electron machines. For DTMs, we can see that only Be-10 is absent from this list of DTMs of the LEP campaign. However, this radionuclide (produced via spallation reaction with Cr, Fe and Ni) has never been above the Declaration Threshold in the campaigns treated till today.

Finally, the radionuclide inventory for the LEP campaigns C001 and C005 was obtained

from direct Fluka Monte Carlo simulations performed in 2013, before the beginning of this thesis [Magistris and Otiougova 2013]. The direct Fluka simulations indicate that the major ETM radionuclides, for stainless steel activated at the LEP, are Ti-44 and Co-60. The predicted DTM radionuclides are H-3, Fe-55 and Ni-63. Direct Fluka calculations and Actiwiz simulations are in agreement. More intensive Actiwiz calculations should however be performed in order to take into consideration a larger number of activation scenarios.

5.2.3 Quantification of ETM's activity

The quantification of the ETM radionuclides is made via systematic γ-spectrometry measurements performed on every waste package. For operational reasons we perform either two or 4 measurements on the waste package according to their weight. In particular, if the waste package has a mass above three tons 4 measurements are systematically performed to reduce the risk of missing radionuclides due to self-attenuation of the waste material.

The description of the measurement system adopted at CERN is made in Ch. 3. We recall here that a normal counting time for waste characterization is fixed at 15 minutes. This duration is enough to reach an MDA which is at least one order of magnitude below the Declaration Thresholds of the radionuclides of interest. For special radionuclides with a very low DT (recall the case of Ag-108m discussed at the beginning of this chapter), either a longer measurement time is set or samples are collected to calculate a scaling factor. If the scaling factor is not applicable, the mean activity method applies.

A typical measurement setting for radioactive waste package measurements at CERN is showed in Fig. 5.15.

The list of ETM radionuclides found in the campaigns of the SHERPA project (as of June 2017) and quantified above the Declaration Threshold is given in Tab. 5.9.

Table 5.9: List of ETM radionuclides quantified above the Declaration Threshold per material type. The table includes radionuclides produced at both proton and electron accelerators.

Material	**ETMs above DT**
Aluminium	Na-22, Ti-44, Co-60, Ag-108m
Copper	Ti-44, Co-60, Ag-108m
Steel	Ti-44, Mn-54, Co-60, Zn-65, Ag-108m, Bi-207

Figure 5.15: Spectrometric measurement system deployed at CERN for the measurement of ETM radionuclides [Bruno et al. 2017].

We should notice here that some radionuclides appear in the table but were quantified above the Declaration Threshold in a very limited number of cases. For instance, Ti-44 appears in both aluminium and copper packages but is above the DT in only two packages per material type. Similarly, Mn-54 is present above the DT in 6 packages of the family steel, Zn-65 in 5 packages of steel and Bi-207 in a single steel package.

The uncertainty quantification introduced in Ch. 4 is not presently implemented at CERN. The specific activities of the radionuclides declared to ANDRA, and used for the calculation of the IRAS, consist of the upper bound of the symmetric interval calculated at a 95% confidence level. In this context the uncertainty associated with the activity of the radionuclides is set to 0 by default.

Among the 8 campaigns included in this summary, three are made of waste from the LEP accelerator. For legacy metallic waste produced at CERN, no difference was found in terms of important ETM radionuclides from hadron and electron machines. Moreover, the activity levels are similar.

The preliminary radionuclide inventory (see Tab. 5.7) includes a large number of radionuclides but only a very few are important for waste characterization. One reason is

that we mainly deal with legacy waste of long decay times (generally above three years). Some radionuclides, such as Mn-54, Co-57 and Zn-65, have an half-life below one year. The calculations for short decay times can predict short-lived radionuclides but at the moment of their measurement they have completely decayed.

The histograms of the specific activity of two major ETM radionuclides (Co-60 and Na-22) in steel and aluminium materials are showed in Fig. 5.16 as an example. From the figure we can see that, with very few exceptions, these specific activities are very low. For instance, the average specific activity of Co-60 on steel packages (we considered here only the packages for which the level of Co-60 is above the DT) is 0.94 Bq/g. The maximum found is 13.83 Bq/g. In the figure, n indicates the number of packages for which the activity of the radionuclide is above the Declaration Threshold (0.1 Bq/g for both Co-60 and Na-22).

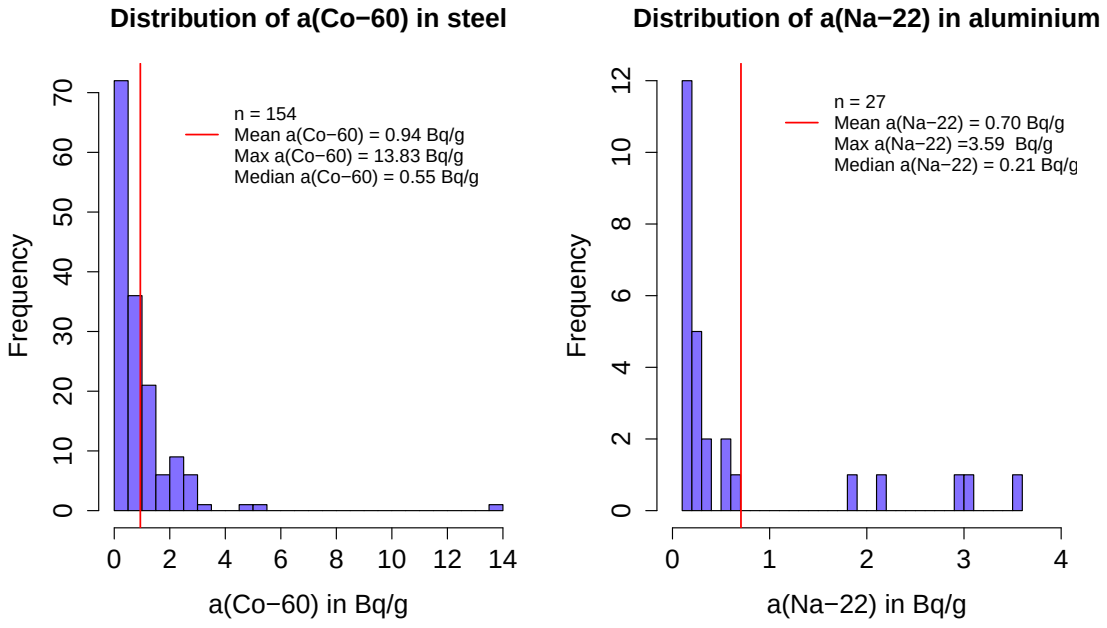


Figure 5.16: Histograms of the specific activities of Co-60 in steel and Na-22 in aluminium. The figure shows only the values above the Declaration Threshold. The red line indicates the average specific activity. Waste produced at both hadron and electron machines are included.

5.2.4 Quantification of the activity of DTMs and ITMs

DTM's activity

The quantification of the DTM radionuclides is made either via the use of scaling factors or via the mean activity method, as discussed in Ch. 3. The key nuclides identified for the SHERPA project are Na-22 in aluminium, Co-60 in copper and Ti-44 and Co-60 in steel.

If a γ-emitter is found systematically on a batch of waste, it is considered for a correlation test with various DTM radionuclides. It is important to notice that, for the majority of radioactive waste treated in the context of the SHERPA project, the decay time is above three years (see Fig. 2.7 of Ch. 2). Therefore, short-lived radionuclides have often completely decayed, leaving only a limited choice for potential key nuclides.

The list of DTM radionuclides, quantified above the Declaration Threshold per material and machine types, is given in Tab. 5.10.

Table 5.10: List of DTM radionuclides quantified above the Declaration Threshold per material and machine types.

Material	**DTMs above DT (hadrons)**	**DTMs above DT (electrons)**
Aluminium	H-3	H-3
Copper	H-3, Ni-63	Ni-63
Steel	H-3, Fe-55	H-3, Fe-55, Ni-63

Some radionuclides appear in the table but were quantified above the Declaration Threshold in a very limited number of cases. For instance, Ni-63 is above the DT in only 4 steel packages.

In terms of DTM radionuclides, a few differences are observed between electrons and protons machines. For instance, H-3 is not a relevant DTM in copper activated at electron machines but is a very common spallation product at hadron accelerators. The radionuclide Ni-63 is relevant in steel activated at the LEP and a typical production channel is the reaction (γ, n) with the element Nickel.

The histogram of the specific activity of H-3 in aluminium waste is showed in Fig. 5.17 as an example. The activity of H-3 is often below 200 Bq/g. The average specific activity of H-3 in the packages for which the level of H-3 is above the DT is 58.74 Bq/g. The

maximum found is 907.20 Bq/g. In the figure, n indicates the number of packages for which the specific activity is above the Declaration Threshold (1 Bq/g).

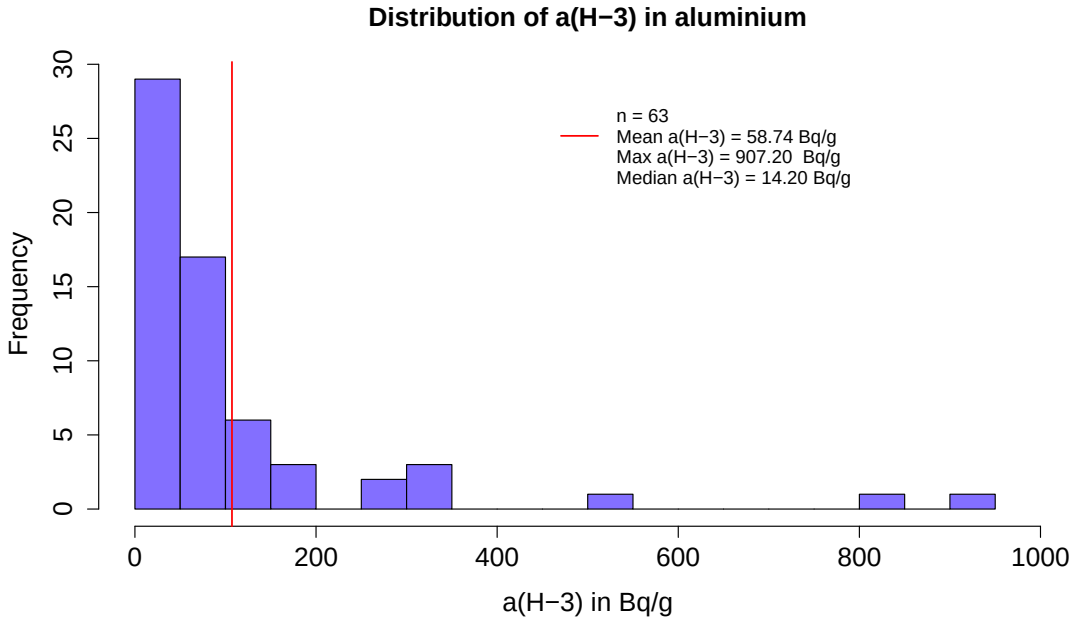


Figure 5.17: Histogram of the specific activity of H-3 in aluminium from both hadron and electron accelerators. The figure shows only the values above the Declaration Threshold. The red line indicates the average specific activity.

In the context of the SHERPA project (as of June 2017), 465 samples have been collected and 854 radiochemical analyses were performed to quantify the specific activity of DTM radionuclides. This corresponds to about one sample collected per 1.5 tons of characterized radioactive waste and 5 radiochemical analyses per 4 tons of waste. Samples were collected using authoritative, random and systematic sampling, as discussed in Sec. 3.2.3 of Ch. 3. Radiochemical analyses and liquid scintillation counting were briefly discussed in Sec. 3.2.4.

Finally, Tab. 5.11 presents mean activities and scaling factors for the SHERPA project. The table indicates the name of the campaign (C00x, where x is a sequential number), the material and the method applied to quantify the activity of the DTMs: MA if the mean activity (in Bq/g) is calculated or SF if the scaling factor is used. Within parenthesis are given either the radionuclide for which the mean activity is calculated, or the pair DTM/KN, if the scaling factor applies. For campaign C006 the letters “a” to “d” indicate

a specific sub-campaign.

Table 5.11: Summary of mean activities and scaling factors for the campaigns of the SHERPA project characterized between 2015 and 2017.

ID	Material	MA in Bq/g	SF
C000	Aluminium	0.75 (H-3)	-
	Steel	0.07 (H-3), 2.17 (Fe-55)	-
C001	Aluminium	62.8 (H-3)	-
	Steel	0.07 (H-3), 0.75 (Fe-55)	-
C002	Aluminium	-	4.43 (H-3/Na-22)
	Steel	43.96 (Fe-55)	32.61 (H-3/Ti-44)
	Copper	-	8.10 (Ni-63/Co-60)
C003	Aluminium	-	1498.62 (H-3/Na-22)
	Steel	0.04 (H-3), 6.54 (Fe-55)	-
C004	Aluminium	-	70.86 (H-3/Na-22)
	Steel	24.16 (Fe-55)	70.48 (H-3/Ti-44)
			14.57 (H-3/Co-60)
	Copper	-	9.22 (Ni-63/Co-60)
C005	Aluminium	62.8 (H-3)	-
	Steel	0.07 (H-3), 0.75 (Fe-55), 1.61 (Ni-63)	-
C006	Steel (a)	0.10 (H-3), 0.99 (Fe-55)	-
	Steel (b)	1.27 (H-3), 0.85 (Fe-55)	-
	Steel (c)	5.56 (H-3), 4.51 (Fe-55)	-
	Steel (d)	0.16 (H-3), 7.45 (Fe-55)	-
C008	Steel	0.13 (H-3), 0.70 (Fe-55), 0.66 (Ni-63)	-

From Tab. 5.11 one can see that if the mean activity method is used, the final specific activity of the DTMs can be below the Declaration Threshold. In the case of the scaling factor, the declaration of a DTM depends on the activity of the key nuclide and can vary from package to package. For the SHERPA project, it was decided that the activities of the radionuclides appearing in the radionuclide inventory are systematically declared to ANDRA and their activities are used to calculate the IRAS. This conservative choice has however a very limited impact (below a few percent) on the final IRAS of a waste package.

By definition, the mean activity method is conservative. This is due to the fact that the arithmetic mean is not a robust average content estimator (especially for asymmetric distributions). Moreover, the upper bound of the 95% confidence interval is presently chosen at CERN for the declaration of the specific activity (and the calculation of the IRAS). Even with such a conservative assumption however, the activity levels of the radioactive waste produced at CERN are well below the activity limits for VLLW.

A last remark should be made about the special case of Ag-108m. For campaign C004,

traces of Ag-108m were detected in 16 packages of various metals. Its specific activity was however very close to the MDA and, often, a clear decision could not be made. We therefore checked for a correlation with the key nuclide Co-60 and found a correlation coefficient $r = 0.66$. We calculated the geometric scaling factor ($\overline{G}_{SF}^{ag108m} = 0.04$) and used it to quantify the activity of Ag-108m for this campaign.

Ag-108m is classified as a DTM or an ETM radionuclide depending on the campaign and the material considered. For this reason the scaling factor calculated for campaign C004 is not showed in the previous table.

ITM's activity

The specific activity of the ITM radionuclides is estimated by multiplying the specific activity of the key nuclide in a package and the geometric correlation factor, which is calculated according to Eq. 2.10 of Ch. 2. As of June 2017, we found that only C-14 (19 packages), Cl-36 (19 packages) and Ar-39 (15 packages) can be quantified above the DT.

The average specific activity of C-14 above the Declaration Threshold is 0.32 Bq/g and the maximum is 1.10 Bq/g. C-14 was found in the aluminium packages of campaign C003 and the key nuclide used to estimate its activity is Na-22.

The average specific activity of Cl-36 above the Declaration Threshold is 0.01 Bq/g and the maximum is 0.03 Bq/g. Cl-36 was found in the steel packages of campaign C003 and the key nuclide used to estimate its activity is Co-60.

Finally, the average specific activity of Ar-39 above the Declaration Threshold is 23.89 Bq/g and the maximum is 54.36 Bq/g. Ar-39 was found in the steel packages of campaign C003 and the key nuclide used to estimate its activity is Co-60.

It should be noted that the calculated specific activity of the ITM radionuclides could overestimates the true activity because a conservative value of the specific activity of the key nuclide (upper boundary at 95% confidence level) is used for the calculations. We can notice however that only very few radionuclides exceeds the Declaration Threshold, and the activities are still very low with respect to the specific limits.

Tabs. 5.12- 5.13 show the summary of the correlation factors for the SHERPA campaigns

described in this section. We can notice that campaigns C005 and C007 are missing. For campaign C005 all the important β-emitters are DTMs and no ITM radionuclide was identified via calculation. Campaign C007 is made of a family of waste (cables) which is outside the perimeter covered by the SHERPA project and is not discussed here.

A comparison between calculated correlation factors and experimental scaling factors is discussed in the last section of the present chapter.

Table 5.12: Summary of correlation factors for aluminium and copper waste characterized between 2015 and 2017.

Material	Campaign	C-14/Na-22	H-3/Co-60	Fe-55/Co-60
Aluminium	C000	$5.3 \cdot 10^{-3}$	-	-
	C002	$5.3 \cdot 10^{-3}$	-	-
	C003	1.81	-	-
	C004	0.04	-	-
Copper	C002	-	8.6	0.96
	C004	-	-	1.63

Table 5.13: Summary of correlation factors for steel waste characterized between 2015 and 2017.

Material	Campaign	C-14/Ti-44	Cl-36/Ti-44	Ar-39/Co-60	Ca-41/Ti-44	Ni-63/Co-60	Cl-36/Co-60	C-14/Co-60	Ca-41/Co-60
Steel	C000	$3.8 \cdot 10^{-3}$	$4 \cdot 10^{-4}$	0.41	$1.6 \cdot 10^{-3}$	-	-	-	-
	C001	-	-	-	-	30	-	-	-
	C002	$3.8 \cdot 10^{-3}$	$4 \cdot 10^{-4}$	0.67	$1.6 \cdot 10^{-3}$	-	-	-	-
	C003	$4.5 \cdot 10^{-3}$	-	16.19	-	-	$9.1 \cdot 10^{-3}$	-	-
	C004	$3.8 \cdot 10^{-3}$	$4 \cdot 10^{-4}$	1.84	$1.6 \cdot 10^{-3}$	-	-	-	-
	C006	$3.8 \cdot 10^{-3}$	$4 \cdot 10^{-4}$	0.29	$1.6 \cdot 10^{-3}$	-	-	-	-
	C008	-	-	0.016	-	-	$2 \cdot 10^{-5}$	$1.6 \cdot 10^{-3}$	$8 \cdot 10^{-5}$

5.2.5 Calculation of IRAS

Using the activities of ETM, DTM and ITM radionuclides described in the previous sections, we can calculate the IRAS of each single waste package according to Eq. 1.11. The summary of the IRAS for the waste population treated (as of June 2017) is given in Fig. 5.18, where the histograms show the distributions of the IRAS per material type.

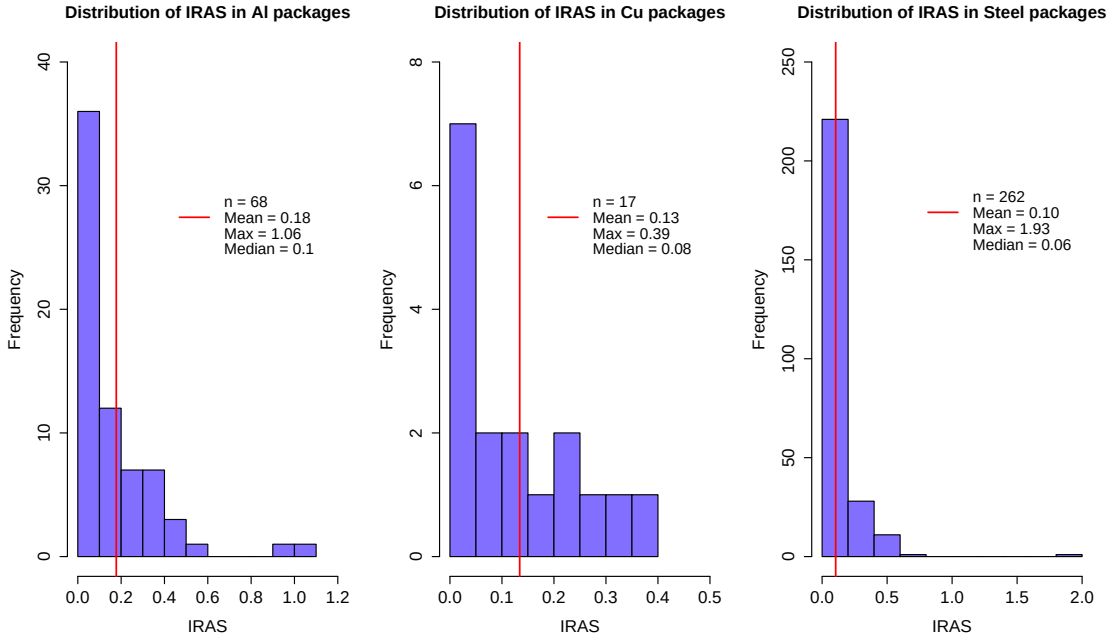


Figure 5.18: Histograms of the IRAS calculated for the three material families of the SHERPA project. The red line indicates the average IRAS.

It is easy to see that the IRAS for the 347 packages considered is well below the limit of 10, fixed by ANDRA for the acceptance of a waste package in the disposal facility. The maximum IRAS found is 1.93 on a steel waste package. If a distinction per material type is not made, the average IRAS over the entire waste population is 0.12, i.e. two orders of magnitude below the limit of 10. The standard error of the mean IRAS is 0.02 (at a 95% confidence level).

The contribution to the IRAS of major radionuclides, per material family, is showed in Fig. 5.19.

Starting from the top we can see the major contributors for aluminium, copper (central

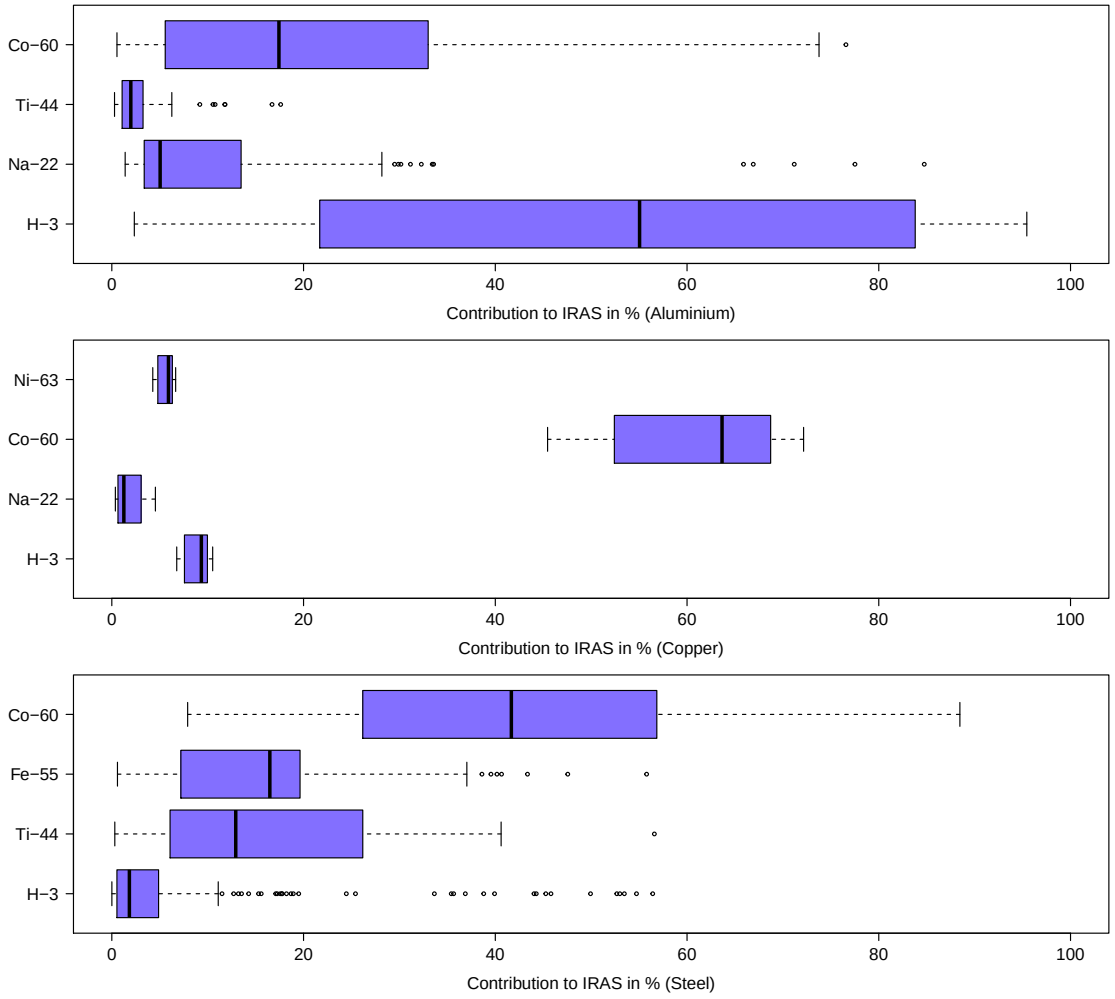


Figure 5.19: Contribution of major radionuclides to the IRAS for the 347 waste packages of the SHERPA project treated before June 2017.

panel) and steel (bottom panel). Depending on the specific waste package the contribution to the IRAS of each radionuclide can vary strongly. For instance, the contribution of H-3 in aluminium waste can span from about 0% up to above 90%.

The central and bottom panels show that in copper and steel radioactive waste the ETM radionuclides (mainly Co-60) are the most important contributors to the IRAS. The experimental results reflect the fact that γ-emitters are very important in the characterization of VLLW, according to their low class (see Sec. 1.3.1 for more details) and must be carefully measured.

In the case of aluminium waste, we found that H-3 plays an important role. This is

related to the difference of half-life between the key nuclide Na-22 ($T_{1/2}$ =∼2.6 years) and the DTM H-3 ($T_{1/2}$ =∼12.3 years). In particular, when the decay time increases the activity of Na-22 decreases more rapidly than the activity of H-3, and the scaling factor increases rapidly as a function of the decay time. A detailed description of a similar scenario is given in Ch. 3 (see Fig. 3.9 and its description).

Of course other radionuclides can be present in specific waste packages but their impact on the IRAS is below 1% and are neglected in Fig. 5.19 for clarity.

As a conclusion, we would like to point out that only a small fraction of the predicted radionuclides is effectively present within the waste population of the SHERPA project treated and characterized so far. It is also important to consider that, among the important radionuclides, only a part of them has an impact on the IRAS in terms of contribution.

According to the definition of the IAEA given in reference [IAEA 2007], laboratories such as CERN produce radioactive waste which is "complex and highly heterogeneous". If we consider the total amount of past, present and future waste this definition seems appropriate. At the same time, we have demonstrated that a waste population with similar material composition and radiological history - as is the case of SHERPA waste - is simple and homogeneous in terms of induced radioactivity. An appropriate definition of the boundaries of a waste population is therefore of outmost importance to achieve a robust radiological characterization.

A discussion in terms of costs, limits and future developments of the characterization strategy proposed in this thesis is given in the last section of the present chapter.

5.3 Discussion

5.3.1 Time and financial aspects of the characterization

In this section we describe some time and financial aspects related to the radiological characterization of very-low-level radioactive waste at CERN. The information given here is based on the feedback from the disposal campaigns of the SHERPA project for the years 2015-2017.

We will primarily focus on the aspects which are related to the radiological characteriza-

tion of legacy waste (such as direct calculation of the radionuclide inventory, measurements of ETMs and analyses of γ-spectrometries, sampling and radiochemical analyses of DTMs and calculation of IRAS), leaving out engineering tasks such as sorting, packaging and transport towards the disposal facility.

One of the very first task of the radiological characterization of a waste population is the calculation of the radionuclide inventory. The calculation lies on the collection of preliminary information about the nature of the waste, such as the elemental composition of the material and its history. Depending on the availability of the information and the size/heterogeneity of the waste population, this task can be usually performed within 2-4 weeks by an engineer or a physicist. This task is not however a full time task. Its accomplishment requires the collaboration with IT personnel to extract data from databases, with technicians to clearly identify the perimeter of the population in the temporary waste storage and to perform preliminary elemental analyses to identify major elements (for an example see Sec.5.1).

A visual inspection of the waste is important because it allows us to identify possible foreign materials or dangerous chemicals that should be separated before treatment. Among dangerous materials that can be found on radioactive waste we can cite some types of asbestos (friable), derived waste from electric and electronic devices (electronic boards and batteries) and metallic lithium. The visual inspection is also important to identify risks and ensure safe maintenance/dismantling of the waste items during waste cutting, compaction and packaging.

At this stage, the engineer/physicist calculates the preliminary correlation factors to identify important radionuclides and discriminate between DTM radionuclides (that will be measured) and ITM radionuclides whose activity will be estimated via calculation.

Once the radionuclide inventory is established, the activity of ETM and DTM radionuclides can be quantified. For ETM radionuclides systematic γ-spectrometry measurements are performed on every waste package. The experience accumulated within the frame of the SHERPA project indicates that 6-8 waste packages per day can be measured by a technician working at full time. The tasks of the technician includes the maintenance and positioning of the packages from the temporary storage to the measurement site, the

actual measurement of the package, the quality control associated with the mechanical and physical characteristics of the waste package and the analyses of the results. Depending on the difficulty of the spectra to analyse (number of radionuclides, interferences and level of activity), the technician can usually perform between 10 and 30 analyses per day.

The technician is also responsible for the measurement and quality control of the background at the measurement site, the calibration of the instruments and the maintenance of the measurement system. With the industrialization of the processes of the SHERPA project the tasks related to γ-spectrometry require a full time technician.

For the quantification of DTM's activity the more demanding tasks are the collection and preparation of samples and the radiochemical analyses. The amount of work required by the collection of samples strongly depends on the nature of the waste material and its physical form. It can be performed during the phase of dismantling of the items and often requires special tools and equipments such as mechanical and electrical saws (that can be cooled), drills, pliers, magnets and protection for a safe maintenance (including masks, gloves and dedicated laboratories to avoid contamination or dispersion of materials). Special care must be taken when the tools used to collect samples generate heat. In fact, some radionuclides (such as H-3) can partly evaporate when the material is heated and their activity cannot be correlated to the waste population from which the sample is withdrawn.

Once the samples are collected and conditioned for the expedition - the amount of material collected is generally below 20 grams - they are sent to external laboratories to perform radiochemical analyses. Based on the experience of the SHERPA project, 2-to-3 months are required to receive the results of the measurements from the date of reception at the external laboratory.

The cost associated with the radiochemical analyses depends on the radionuclide to be measured. For H-3, Cl-36, Fe-55 and Ni-63 the cost of the analyses varies between 200 to 300 CHF per sample. Often, charges for sample preparation and project management tasks must be added.

One last task of the characterization process is the calculation of the IRAS and the quality control associated with the consistency of the data. This task is usually performed

by the project manager or by the scientist in charge of the calculations. This task is performed in 1-to-2 weeks, including identification of possible mistakes, their correction and analysis of sources of error.

In conclusion, the experience of the SHERPA project suggests that 6 months are necessary, on average, to complete the characterization of a campaign of waste. Of course this duration can sensibly change depending on the size of the waste population and the availability of a laboratory to perform radiochemical analyses.

5.3.2 Limits, ongoing improvements and future work

In this section we discuss risks and limitations of the characterization method. We describe the improvements that were identified after the first test and implemented into operational methodologies, as well as recommendations on future research axes.

The discussion will follow the main stages of the characterization process as showed in Fig. 1.10 of Ch. 1 (for a succinct description of the process see the summary of Ch. 1).

Identification of a campaign of radioactive waste

When a campaign of radioactive waste of the SHERPA project is identified, the main materials (aluminium, copper and steel) are separated and conditioned in different packages. The sorting stage is very heavy both in terms of time and cost and can represent an increased risk of exposure for the workers. At CERN a new procedure was developed in 2016-2017 to characterize waste without a preliminary sorting phase. According to this new method the materials can be mixed in the final waste packages (the cutting and compaction operations can be made without previous dismantling and sorting of the metals) [Magistris et al. 2017b].

This new procedure, based on conservative assumptions about the major radionuclides contributing to the IRAS and dose rate measurements, was submitted at ANDRA in 2017 and accepted for the characterization of future radioactive waste produced at CERN. It is foreseen for this procedure to become operational for all metallic and non-metallic radioactive waste between 2017 and 2018.

Until now the characterization strategy is applied to legacy waste with more than three years of decay time. For new waste it is reasonable to assume that the decay time will be shorter and a higher number of radionuclides (with lower half-life) will be present. We are convinced that the methodology presented in this thesis can also be extended to future waste, and are curious to test its robustness on new classes of waste. This belief is based on the fact that ETM radionuclides will be systematically measured on every waste package. Moreover, the extensive numerical experiments performed suggest that the DTM important for VLLW characterization are few.

In this thesis we have focused on the radiological characterization of metallic waste. However, the same method was adapted to the characterization of 1000 m^3 of cables made of plastic and metals (see Sec. 5.1) and the characterization of air filters.

Calculation of the radionuclide inventory, correlation factors and ITMs

The calculation of the preliminary radionuclide inventory is based on the information collected before a disposal campaign starts. As we have discussed extensively in the previous chapters, we often consider a large number of activation scenarios allowing us to estimate a long list of radionuclides that can be produced by activation at CERN.

The calculation of the radionuclide inventory can be affected by various limitations including the consideration of non-representative scenarios, the presence of unexpected trace elements on the elemental composition used and the wrong classification of DTM and ITM radionuclides.

In Ch. 3 we have seen that among all the input parameters of the vector scenario the decay time is the one that has the highest impact on the determination of correlation factors. Using values of decay times non adapted to a specific batch of waste can have a big impact on the estimation of the activities of DTM and ITM. It is important to recall however that major pure or quasi-pure β-emitters and low-energy X-ray emitters are quantified trough radiochemical analyses. If the experimental data does not agree with the calculations we question the realizations of the scenario considered. At the end, for major DTMs, only experimental data is used for the quantification of the specific activity. An example of difference between calculation and experimental data is discussed later in this

section.

The second limitation that we pointed out is the presence of unexpected trace elements or impurities. The risk associated with this event is low because we consider by default a large number of impurities on the chemical compositions used for the simulations.

The creation of random chemical compositions from elements distribution has been tested on copper (see Ch.2). In a near future we plan to extend such a method also to other metals and materials. As an example, we want to compare here the radionuclide inventory of the cathodic copper with random variations discussed in Sec. 2.2 and a standard high-purity copper composition commonly used at CERN for the calculations performed in the context of the SHERPA project. The second copper grade is called CuOFE (high-purity oxygen-free copper for electronic applications). The list of the trace elements of cathodic copper is given in Tab. 2.6. All the trace elements of copper CuOFE, with the exception of the element Hg, are also present in cathodic copper. The radionuclides contributing for more than 1% to the IRAS are Co-60, Ni-63 and H-3 for both chemical compositions. The contributions to the IRAS are showed in Tab. 5.14.

Table 5.14: Comparison of major radionuclides obtained with cathodic copper and copper CuOFE. C_{IRAS} is the contribution of a given radionuclide to the IRAS.

Radionuclide	**C_{IRAS} cathodic Cu**	**C_{IRAS} CuOFE**
Co-60	93.91%	96.58%
Ni-63	3.17%	1.5%
H-3	2.27%	1.6%

From the table one can see that the variation of the amount of trace elements in the elemental composition has a very limited impact on the contribution of major radionuclides to the IRAS. We performed this simple test also for other grades of copper and the results are similar to the one obtained for CuOFE. Of course, this test is valid only for copper and cannot be generalized to other materials such as steel. This is due to the fact that the material copper is generally highly pure and the amount of trace elements is controlled (see Ch. 2 for a detailed discussion). The methodology of random extraction of the amount of trace elements from probability distributions will be implemented at CERN also for other families of materials in a near future.

Another identified risk associated with the calculation of the radionuclide inventory is the possible wrong classification of DTM and ITM radionuclides. This classification is based on the contribution of each single radionuclide to the IRAS, as discussed in Sec. 2.1.3. As we have detailed in previous chapters, if a radionuclide contributes above a pre-defined threshold to the IRAS, its activity is evaluated from analyses performed on samples. If not, its activity is evaluated from calculation. The impact of the wrong classification (DTM vs ITM) of a radionuclide or its activity evaluation is minimal because the two method used (scaling factor or mean activity and correlation factor) give consistent results, as demonstrated in [Zaffora et al. 2017a]. In particular we showed that the confidence intervals of the theoretical correlation factors include the experimental scaling factors in the majority of cases.

Tab. 5.15 shows the comparison between CFs and SFs for all the campaigns of the SHERPA project where scaling factors apply.

Table 5.15: Comparison of correlation factors and scaling factors for the disposal waste campaigns 2 to 4. The lower and upper boundaries in parentheses are given within a 95% confidence level.

Material	**KN**	**DTM**	**CF**	**SF C002**	**SF C003**	**SF C004**
Aluminium	Na-22	H-3	11 (0.07, 1500)	4.4 (0.3, 65)	1500 (66, 34100)	71 (9, 570)
Copper	Co-60	H-3	1.7 (0.1, 22)	-	-	15 (0.9, 250)
		Ni-63	1.6 (0.07, 34)	8.1 (1.3, 5.2)	-	9.2 (0.5, 190)
Steel	Ti-44	H-3	84 (4, 1700)	33 (0.3, 3350)	-	71 (9, 530)

In the previous table we can see that the confidence intervals of CFs systematically contain experimental SFs with the exception of the scaling factor of H-3 in aluminium for campaign C003. In this case the experimental SF is the same as the upper bound of the CF. The calculated geometric CF is in general a good estimator of the geometric SF and CFs predict scaling factors within 1 order of magnitude.

After the identification of this discrepancy in the campaign C003 we collected complementary information. We found that the waste was stored for at least 20 years in CERN's storage, instead of the reference range from 3 to 30 years assumed in the calculations. We performed a new set of simulations fixing all the input parameters with the exception of the decay time that now spans from 20 up to 40 years. The new calculation includes 1050

scenarios. The new geometric CF is 280 and the confidence interval is (11, 6900). The amount of Na-22 ($T_{1/2}$ = 2.6 years) decreases faster than H-3 ($T_{1/2}$ = 12.3 years) and for long decay times the ratio of activities H-3/Na-22 increases (see Sec. 3.3 for more details). The confidence interval now includes the experimental SF and the calculated CF is within one order of magnitude of the SF. Another set of calculation was performed including only decay times longer or equal to 30 years. The new CF is now 800 and the confidence interval (99, 6500). As expected, the value of CF increases again. Such kind of test can also be used to identify a plausible decay time for legacy waste of unknown radiological history.

In the previous example we showed that calculations and experiments are in agreement after that complementary information was collected. For legacy waste the radiological history of the waste is often unknown and large scale calculations can provide a robust estimation of scaling factors. We also showed how new input data can be used to modify the analytical calculations in order to adapt the predictive model to experimental results.

As discussed in the case study at the beginning of this chapter, we also tested experimentally the importance of ITM radionuclides. To do so we collected samples from a waste population of shredded copper and measured via radiochemical analyses the most important ITM radionuclides. We have seen that only very few samples are above the MDA and their measurement is not justified from both a technical and financial point of view.

We would like to conclude this section stating that, the results obtained in the context of the SHERPA project are in agreement with the theoretical activation studies performed over the last years at CERN (for comparison, see [Mitaroff and Silari 2002][Brugger et al. 2006]). We believe that the quality control put into place at CERN can identify anomalies (if any) that could be present in spite of the amount of calculations and measurements performed.

Quantification of ETMs and the key nuclide

The systematic measure of ETM radionuclides is very time consuming. In the large majority of the packages of the SHERPA project only a very limited number of γ-emitters were identified. Whenever the contribution of the γ-emitters to the total counting is fairly

constant the direct spectrometry measurement can be replaced by a dose rate measurement and a conversion factor between dose and activity. The quantification of the activity of ETM radionuclides via dose-rate measurements and conversion factors is one of the methodology suggested by the IAEA to characterize low-level-waste and is described in [IAEA 2007]. This method is commonly employed at nuclear power plants for waste streams with constant nuclide vectors and is presently under consideration at CERN for the characterization of future waste.

The selection of the key nuclide is important for a robust characterization of DTM and ITM radionuclides. For the time being, a limited number of validated ETM are identified as key nuclides (Na-22 in aluminium, Co-60 in copper and Ti-44 and Co-60 in steel). A systematic evaluation of the correlation between each measured ETM and DTM radionuclides is also under consideration. This systematic approach would be useful to identify the most appropriate key nuclide.

Quantification of DTMs, mean activity and scaling factors

A number of considerations must be made about the quantification of the activity of DTM radionuclides. These considerations are related to the sampling process and the number of samples, the evaluation of scaling factors per campaign, the correction of the scaling factors per decay time, the use of different average content estimators and the use of linear models when estimating the robustness of the correlation between ETM and DTM radionuclides.

We would like to point out that the case study discussed at the beginning of this chapter is a quasi-ideal situation in which the analyst has access to the entire population and can collect census data. In reality such data are rarely available and often for logistic/maintenance reasons some parts of the waste population are not accessible.

In such a case the sampling of a waste population can be authoritative (see Sec. 3.2.3). Moreover, the number of samples is limited by cost/time constraints of the project. In the past years a compromise was made between probabilistic and authoritative sampling because samples were collected systematically from the accessible waste in a container at the hotspot and a sample was taken per type of material. This hybrid sampling was

replaced in 2016 by systematic and random sampling.

After two years of operation of the SHERPA project and the large amount of samples collected we have created a large database of experimental scaling factor that are immediately usable. In the near future we foresee to calculate global scaling factors over an entire waste population - such as all the campaigns of the SHERPA project - using the analyses performed in the past years. This is motivated by the fact that the waste population is low-heterogeneous, the activation mechanisms are similar and the calculated scaling factors, with a single exception, are within one order of magnitude.

If global scaling factors can be applied we will introduce a simple periodic quality control to check for the validity of the scaling factor. This quality control would consists in taking limited samples for a new waste population, measuring the activity of the DTM radionuclides and compare it with the values obtained using the global scaling factor. Of course, a correction of the scaling factor with the decay time should be introduced in order to follow the variation of SF, as discussed in Sec. 3.3.

A larger sample population would also reduce the effect of small populations. In fact, we have seen that a very limited number of samples (< 10) can generate artificially high correlations (r) between the activities of DTM and ETM radionuclides. In this last case linear models and the statistics associated with linear models (such as p-values, F-statistics and R^2) are better qualified to study linear relationships with respect to the Pearson coefficient.

A remark should be made also about the use of the arithmetic mean in the mean activity method. The arithmetic mean is not a robust central tendency estimator for skewed distributions. The median or the geometric mean are not impacted by extreme values and should be considered in future as valid replacement of the mean.

We would like to conclude this section stating that statistical learning methods, including multi-linear regression and regression trees, have been extremely helpful on identifying the parameters that have a strong impact on scaling and correlation factors. We know today that the decay time and, in a minor extent, the irradiation time are the key parameters affecting the behaviour of these ratio estimators.

Uncertainty estimation

When discussing about uncertainty two different approaches are often considered in radioactive waste characterization: "representative" and "conservative".

The representative approach is based on a rigorous application of the rules of uncertainty quantification as described in Ch. 4 and given in the GUM and its supplements [JCGM 2008a] [JCGM 2008b] [JCGM 2011].

This rigorous approach is sometimes very difficult to apply because information can be missing, the real relationship between variables can be unknown or the collection of information is too long and expensive to be practically feasible.

On the other side, authorities encourage waste producers to declare "reasonably conservative" activities and IRAS values, especially when dealing with legacy waste of unknown history.

For this reason, a conservative approach is presently practised at CERN when characterising low-level-radioactive waste. This approach consists of replacing the best estimator of the true activity (like the average activity of an ETM radionuclide in a waste package) by the upper bound of the confidence interval at 95%.

Similarly when the mean activity method applies, the true activity of the DTM radionuclide is estimated by the average plus twice the standard error of the mean. Finally, if scaling factors are used, the true activity of the DTM is estimated using the upper boundary of the confidence level (at 95%) of the activity of KN.

Another aspect of the conservative approach relies on the fact that, even if various radionuclides have an activity below the Declaration Threshold, their activity is used to calculate the IRAS.

As a results of the overestimation of activities the true IRAS of each waste package can be overestimated of about 20%, which corresponds roughly to the uncertainty of the major γ-emitters at a 95% confidence level.

The practice of taking upper boundaries for the activities includes, as a corollary, that the uncertainty of the estimated quantity is set to 0 by default. For this reason the activities

and the IRAS of waste packages are presently declared without uncertainties.

The present situation, even if accepted or even encouraged by the authorities, represents a biased evaluation of the true quantities of interest. In Ch. 4 we showed in detail how the uncertainty can be propagated over activities and the IRAS of both packages and batches.

In the near future we would like to apply the representative approach to the routine characterization of waste in order to propose a more robust estimation of the quantities of interest and their uncertainties. We believe also that a representative uncertainty quantification is necessary to avoid unnecessary burdens to the centre for waste disposal. What we mean is that biased conservative estimation of activities generate "administrative" activities that do not necessarily correspond to the physical reality of the waste. A representative approach should therefore be considered to avoid a faster saturation of the disposal sites.

At CERN, in order to move in this direction, we are creating an uncertainty budget system to systematically quantify the uncertainty of the activity of each radionuclide and method of quantification (direct measurement, scaling factors, mean activity, correlation factors and linear models among others). A first comprehensive application of the uncertainty quantification process was described in Ch. 4 and the case study of Sec. 5.1.

Summary

In this chapter we discussed a case study presenting in detail the application of the new characterization method, we described the industrialization of the characterization method within the frame of the so-called SHERPA project and concluded with the limits that should be addressed and the improvements and future works that should be performed to optimize the characterization process.

The case study represents an ideal case of radiological characterization because we had access to census data, the population was completely accessible and a large number of measurements was performed.

We used the case study to apply for the first time a rigorous uncertainty quantification process to estimate the uncertainties of the activities and the IRAS according to the indication given by the GUM [JCGM 2008a].

The SHERPA project, discussed in the second section of this chapter, is the largest disposal campaign performed at CERN since its foundation. The radiological characterization strategy of this family of waste was developed in the context of this thesis with the collaboration of the personnel of the radioactive waste section of the radiation protection group of CERN. As of June 2017 above 690 tons of radioactive waste have been characterized or disposed of in final repositories without any non-conformity identified by the competent authorities.

Finally, a number of recommendations of optimization and improvement of the characterization method are identified. The discussion about their technical implementation is the topic of the last section of this chapter.

Conclusion

It is far better to foresee even without certainty than not to foresee at all.

Henri Poincaré
The Foundation of Science.

I would rather have questions that can't be answered than answers that can't be questioned.

Attributed to Richard Feynman

The aim of this dissertation has been to propose an engineering solution to radiologically characterize metallic very-low-level radioactive waste (VLLW) produced at particle accelerators. The method is based on a combination of simulations and measurements allowing us to estimate a preliminary radionuclide inventory, to quantify the activities of the major radionuclides and to evaluate the acceptance of a waste population at the disposal facility based on the calculation of an hazard factor called IRAS. The characterization is centred on legacy waste temporary stored at CERN.

This chapter has three sections. The first section is an overview. The second section presents the achievements and the last section illustrates the directions for future work.

Overview

This thesis consists of 5 chapters. Ch. 1 provided an overview of CERN installations and the activation mechanisms at the origin of the production of radioactive waste. We discussed the disposal pathways available at CERN and introduced the overall scheme to characterize radioactive VLLW. This scheme was obtained adapting to the needs of CERN

the characterization strategies given in the international standards and IAEA technical guides of references [IAEA 2007] [ISO 2007] [ISO 2013] [IAEA 2009b]. Standards and guidelines are conceived for waste produced at nuclear power plants. For this reason we needed to consider the specificity of CERN's radiological environments when building the new characterization process.

Ch. 2 described the simulations performed at the beginning of the characterization process to estimate the preliminary radionuclide inventory. The simulations are performed using either analytical or Monte Carlo techniques and require the knowledge of the chemical composition of the waste, the location within the accelrator's tunnel, the energy of the accelerator and the irradiation and decay times. The output of the simulations is a list of radionuclides that can be potentially produced if the materials were activated at common CERN's activation scenarios. In conformity with the vocabulary introduced by the IAEA in [IAEA 2007] the radionuclides can be either easy-to-measure (ETM) from outside a waste package (γ-emitters), difficult-to-measure (DTM) (such as pure-α and pure-β emitters) and impossible-to-measure (ITM) if their experimental quantification is not feasible within reasonable time and cost. In this chapter we introduced also the so-called correlation factors used to evaluate the theoretical relationship between ETM and DTM or ITM radionuclides.

The third chapter provided an overview of the methods to quantify the activity of ETM, DTM and ITM radionuclides. In particular we described the spectrometric measurements performed to quantify the activity of the γ-emitters, the so-called scaling factor (SF) and mean activity methods to evaluate the activity of DTM radionuclides and the correlation factor method to quantify ITM's activity. The scaling factor method requires the collection of a representative sample from the waste population to be characterized. The activity of the DTM's nuclide of interest is measured in each sample together with the activity of the major γ-emitter called key nuclide (KN). If a consistent and repeated correlation between KN and a DTM can be established, a scaling factor can be calculated using linear models, geometric mean of the ratios or other central tendency ratio estimators. The model so found is used to estimate the activity of the DTM radionuclide from the activity of the KN using a scaling factor. When a correlation is not found or the model is not robust enough, the mean activity method is applied instead. The mean activity method consists of

calculating the average activity (or median or geometric mean) over the collected samples. The value so found is applied to each waste package.

Ch. 4 dealt with the quantification of the uncertainty of the activities of the radionuclides found in the radionuclide inventory and the uncertainty of the IRAS, which is the factor used to estimate the acceptance of a batch of waste in the disposal facility. We described in detail the application of the classical GUM approach [JCGM 2008a] [JCGM 2008b][JCGM 2011] to uncertainty quantification and introduced a Monte Carlo based technique to estimate the uncertainty in those cases where an analytical close formulation cannot be used. We found that the major γ-emitters are often the biggest contributors to the uncertainty of the IRAS and that their uncertainty (relative standard deviation) is of the order of 20% for a 95% confidence level.

Finally, Ch. 5 described a complete case study of legacy activated copper from shredded cables, summarized the results obtained via the industrialization of the characterization method in the context of the so-called SHERPA project and discussed the limits of the work performed, the ongoing improvements and the future axes of research. The case study illustrated an ideal characterization case in which census data can be collected from the waste population. The case was also used to quantify for the first time at CERN the uncertainty associated with the IRAS. The SHERPA project, discussed in the second section of the chapter, represents a practical application of the radiological characterization method proposed by the work of this thesis. At the time of writing (June 2017) above 690 tons of legacy, metallic, radioactive waste have been characterized and partially disposed of. The characterization process is accepted by ANDRA and a number of techniques introduced for the first time in this thesis have been used as a model for other families of waste characterized at CERN. In the last section of Ch. 5 we discussed some limits of the characterization method including the high number of measurements to be performed on both waste packages and samples, the need to estimate scaling factors per batch of waste and the risk associated with a wrong classification of DTM and ITM radionuclides. A number of solutions to address such limitations were already implemented or are presently under discussion at CERN. We can cite for instance the calculation of global scaling factors using the experimental data obtained during last 2 years (above 800 radiochemical analyses

performed), the use of dose rate measurements to replace γ-spectrometry and the possibility to mix metal within a waste package to reduce the time associated with the sorting of waste per material family.

Achievements

The discussion that follows illustrates the achievements of this thesis and its impact on the operations related to the disposal of VLLW at CERN.

A major contribution of the present thesis is the design and implementation of a new process to characterize VLLW which is nowadays routinely applied for the radiological characterization of metallic waste. This new characterization method was accepted by ANDRA, is robust (i.e., is applicable to the majority of metallic waste produced at CERN) and can be easily adapted for the characterization of other families of waste, including radioactive cables.

This work has also contributed to change the scope of waste characterization at CERN. In fact, over the last few years we moved from the characterization of individual items of waste with well known radiological history (such as targets for experiments) to the characterization of hundreds of tons of mixed waste with limited or unknown history.

The analytical method developed to establish the preliminary radionuclide inventory made a detailed study of elemental composition of materials unnecessary. This change of approach represents a considerable gain in time and costs. The elemental composition can vary considerably over hundreds tons of waste but the analysis performed demonstrated that elemental probability distributions and average compositions can be used for the successful characterization of large amount of materials, where the specificities of individual items of waste do not affect the properties of the entire batch. The completeness of the calculated radionuclide inventories is today confirmed by thousands of γ-ray spectrometries and more than 800 radiochemical analyses performed on metals.

The scaling factor (SF) method, developed for the characterization of waste produced in nuclear power plants, was tested and modified to take into account CERN's operational constraints. Based on a comprehensive analysis of the ways the SF method is applied

by dozen different countries, we developed a large-scale simulation system to test the applicability of SF to the specificity of particle accelerators. The validity of the SF method was proved both analytically and experimentally and it is today routinely implemented for the characterization of VLLW at CERN.

In 2014 one of the needs of CERN was the creation of an efficient disposal flow of waste towards the appropriate disposal facilities. The definition of SF per batch of waste (instead of calculating global scaling factors, which would have required a large initial time investment) was an efficient technique to start the characterization, and at the same time start collecting data for statistical analysis of CERN waste. The creation of batches is also useful to immediately identify mistakes in the sorting process and solve them efficiently. For instance, if a foreign metal is included in a relatively small batch, its presence will be rapidly identified when the packages are measured and an immediate action can be undertaken, before the error can affect a larger number of packages.

We believe that one of the strengths of the method proposed is its capacity to concentrate the characterization effort on important radionuclides. This is done via the definition of specific parameters, such as the contribution to the IRAS, that are used to generate short lists of radionuclides that must be quantified precisely. The system is created in such a way that negligible radionuclides are discarded from the very beginning of the characterization process, and the most expensive and time consuming measurements are left for the dominant radionuclides.

The radiological history of legacy waste is often unknown. Large-scale calculations validated by systematic γ-spectrometries and radiochemical analyses demonstrated that the method is robust and valid to characterize waste in such an uncertain domain.

The method proposed for the characterization is very efficient because various grades and types of metals were grouped into three, easily recognizable, families: aluminium, copper and steel. The simplification of the families is very useful for the operations related to sorting and packaging. Mistakes can often be identified via a simple visual inspection.

The characterization method is robust with respect to the quantification of difficult-to-measure radionuclides. In fact, the sampling stage is made before the waste treatment

starts. This is possible thanks to the standard radionuclide inventories calculated for metals according to the principles enumerated in this thesis.

We would like to conclude this section recalling that the characterization method proposed in this thesis is today routinely implement at CERN at an industrial scale. Its use has permitted the disposal or characterization of more than 690 tons of metallic VLLW between 2015 and 2017 and the scaling factor and mean activity methods can be adapted to the characterization of new families of waste.

Future axes of research

A number of limits were identified in the characterization method as it is applied today at CERN and need to be addressed to both simplify and optimize the process.

A first goal for future work is the extension of the characterization method to cover new materials and activity ranges. It is foreseen that in 2017-2018 CERN will start a study for the characterization of low- and intermediate-level waste. We believe that the work presented in this thesis represents a good basis upon which the future method can be built.

After two and half years of characterization of metallic waste we have created a consistent data base of scaling factors. In the near future, we will study and test global scaling factors spanning over an entire waste family. Global scaling factors, if applicable, will consistently reduce the costs and time required for characterization because the number of samples to be collected and analysed will be reduced to a minimum. New data will be collected for quality control purposes and to extend the validity range of the scaling factors.

The large number of measurements performed convinced us that the ETM radionuclides present in metallic VLLW is limited and their activity spans over a narrow range. In such a context, the systematic measurement via γ-spectrometry could be reduced to a minimum and/or replaced by dose rate measurements. Dose rate measurements combined with conversion factors between activity and dose are presently under consideration at CERN.

For the time being, VLLW metals are sorted and packaged according to the three families described before (aluminium, copper and steel). The sorting stage is particularly demanding

because every single item must be handled and dismantled. A new methodology presently under discussion proposes to package together different metals. Such a methodology, if implemented, will be based on conservative definitions of radionuclide inventories and would represent a considerable gain in terms of waste processing time.

In this thesis we showed how the uncertainty of activity and IRAS can be quantified for waste characterization purposes. A discussion is presently ongoing at CERN to decide how such a study can be implemented into the operations of waste characterization.

In conclusion, we believe that the characterization of radioactive waste is a complex endeavour requiring a large panel of cross-competencies going from the knowledge of the interactions of particles with matter to the use of advanced calculation tools. When dealing with legacy waste, the complexity is increased by the limited information available. In this context both statistical and numerical methods are a robust and efficient tool to estimate radionuclide inventories and orientate the characterization effort. At CERN these tools are today the foundation of the characterization program. We are convinced that the work described in this thesis can be used for the characterization of future waste produced at CERN but also for waste generated in particle accelerators and research centres where the activation mechanisms are comparable to the ones occurring at CERN.

Bibliography

Ordonnance sur la radioprotection (ORaP) du 22 juin 1994 (Etat le 1er janvier 2014). 814.501, 1994.

Directive IFSN-B04/f en collaboration avec l'Office fédéral se la santé publique (OFSP) et de la Caisse nationale suisse d'assurance en cas d'accidents (Suva). Mesurage de libération de matériaux et de secteurs de zone contrôlées. Version août 2009, 2009.

Accord entre le Conseil Fédéral Suisse, le Gouvernement de la République Française, et l'Organisation Européenne pour la Recherche Nucléaire relatif à la Protection contre les rayonnements ionisants et à la Sûreté des Installations de l'Organisation Européenne pour la Recherche Nucléaire, conclu le 15 novembre 2010, entré en vigueur par échange de notes le 16 septembre 2011. RS 0.814.592.2, Recueil Officiel n. 34 du 23 août 2011, page 3825. Décret n. 2011-1024 du 24 août 2011, JORF n. 0199 du 28 août 2011 page 14594., 2011.

J. Aitchison. On the distribution of a positive random variable having discrete mass at the origin. *J. Am. Stat. Assoc.*, 50(271):901–908, 1955.

Y. Algoet, L. Bruno, M. Magistris, R. Michaud, B. Zaffora, and F.P. La Torre. SHERPA (SHEaR Process Assessment). Project proposal. Technical Report EDMS No: 1428178, CERN, 2015.

ANDRA. Critères radiologiques d'acceptation des déchets TFA. SUR.SP.AMES.02.0007, 2013a.

ANDRA. Spécification d'acceptation de déchets TFA du point de vue de leurs caractéristiques physico-chimiques. SUR.SP.AMES.02.016., 2013b.

I. C. Arsene, L. Bravina, A. B. Kaidalov, K. Tywoniuk, and E. Zabrodin. Energy dependence of nuclear effects in hadron-nucleus collisions. *ArXiv e-prints*, 2007.

ASTM. ASTM D6311 - 98. Standard guide for generation of environmental data related to waste management activities: selection and optimization of sampling design, 2014.

ASTM. ASTM B115. Standard Specification for Electrolytic Copper Cathode, 2016.

M. Avriel. *Nonlinear Programming: Analysis and Methods*. Dover Publications, 2003.

M.M. Barbier. *Induced radioactivity*. North-Holland Pub. Co., 1969.

H. Bateman. The solution of a system of differential equations occurring in the theory of radioactive transformations. *Proc. Camb. Philos. Soc.*, 16:423–427, 1910.

G. Battistoni, F. Cerutti, A. Fassó, A. Ferrari, S. Muraro, J. Ranft, S. Roesler, and P.R. Sala. The FLUKA code: Description and benchmarking. In *Hadronic Shower Simulation Workshop 2006, Fermilab 6-8 September 2006, M. Albrow, R. Raja eds., AIP Conference Proceeding 896, 31-49*, 2006.

C. Bishop. *Pattern recognition and machine learning*. Springer-Verlag, 2006.

C.G. Broyden. The Convergence of a Class of Double-rank Minimization Algorithms - 1. General Considerations. *IMA J. Appl. Math.*, 6(1):76–90, 1970.

R. Bruce, R.W. Assmann, V. Boccone, C. Bracco, M. Brugger, M. Cauchi, F. Cerutti, D. Deboy, A. Ferrari, and L. Lari et al. Simulation and measurements of beam loss patterns at the CERN Large Hadron Collider. *Phys. Rev. ST Accel. Beams*, 17(081004), 2014.

M. Brugger, A. Ferrari, S. Roesler, and L. Ulrici. Validation of the FLUKA Monte Carlo code for predicting induced radioactivity at high-energy accelerators. *Nucl. Instr. and Meth. A*, 562:814–818, 2006.

M. Brugger, A. Ferrari, S. Roesler, and P.R. Sala. Calculation of radionuclide production cross sections with FLUKA and their application in high energy hadron collider studies. In *International Conference on Nuclear Data for Science and Technology*, 2007.

L. Bruno, Y. Algoet, L. Ulrici, M. Magistris, D. Forkel-Wirth, L. Thimonier, and S. Braun. The commissioning of a new treatment centre for radioactive waste from high-energy accelerators at CERN. Technical Report KONTEC 2017, EDMS No: 1745027 v.0.3, CERN, 2017.

BS. BS EN 1978:1998. Copper and copper alloys. Copper cathodes, 1998.

C.P. Camurri, C. Carrasco, R. Leite, and J. Dille. Influence of impurities in cathodic copper on the ductility of copper wires. *J. of Materi. Eng. and Perform.*, 21(7):1474–1478, 2011.

CANBERRA. *Genie 2000 v.3.2. Customization tools manual*, 2009.

CANBERRA. Technical advantages of ISOCS^{TM}/LabSOCS^{TM}. Technical report, 2012.

CDA. Copper and Copper Alloys. Compositions, Applications and Properties. Copper Development Association, Publication No 120. Technical report, 2004.

CEN. CEN - EN13347. Copper and copper alloys rod and wire for welding and braze welding, 2002.

CERN. The Accelerator Complex. `http://home.web.cern.ch/about/accelerators`, 2015.

A. Chabrol and S. Girois. Stabilisation du PVC. *Techniques de l'ingénieur*, Base documentaire: TIB138DUO(AM3233 V2), 2013.

A. W. Chao, K.H. Mess, M. Tigner, and F. Zimmermann. *Handbook of Accelerator Physics and Engineering*. World Scientific, 2013.

W.G. Cochran. *Sampling techniques*. John Wiley and Sons, 1977.

C.W. Coghran and P.P. Egeghy. Methods of dealing with values below the limit of detection using SAS. In *Southeast SAS User Group*, 2003.

L.A. Currie. Limits for qualitative detection and quantitative determination. *Anal. Chem.*, 40(3):586–593, 1968.

L.E. De Geer. Currie detection limits in gamma-ray spectroscopy. *Appl. Rad. Isot.*, 61 (2-3):151–160, 2004.

K. Delbeke and P.H. Rodriguez. Copper Concentrates. Environmental and Human Health hazard classification. Application to the assessment of substances "Harmful to the Marine environment" (HME) as set out under the 2012 Annex V MARPOL Convention amendments. Technical report, 2014.

C. Duchemin, M. Magistris, F. Pozzi, and M. Silari. Clearance of LEP superconducting radiofrequency acceleration system. report submitted to the Swiss Federal Office of Public Health (OFSP). Technical Report CERN-RP-2016-211-REPORTS-TN v.4, EDMS No: 1744306 v.4, CERN, 2017.

W.L. Dunn and J.K. Shultis. *Exploring Monte Carlo methods.* Elsevier Science, 2011.

B. Efron and R.J. Tibshirani. *An introduction to the bootstrap.* Chapman and Hall, 1993.

EPA. Guidance for the data quality objectives process. EPA QA/G-4. Technical report, Office of environmental information, U.S. Environmental Protection Agency, 1994.

EPA. Guidance for data quality assessment. Practical methods for data analysis. EPA QA/G-9. Technical report, 2000.

EPA. RCRA Waste Sampling Draft Technical Guidance. Planning, Implementation and Assessment. EPA530-D-02-002. Technical report, Office of solid waste, U.S. Environmental Protection Agency, 2002.

J.J. Faraway. *Linear models with R.* Chapman and Hall, 2004.

A. Ferrari, P.R. Sala, A. Fassó, and J. Ranft. Fluka: a multi-particle transport code. Technical report, CERN, 2005. CERN-2005-10 (2005), INFN/TC-05/11, SLAC-R-773.

R. Fletcher. A new approach to variable metric algorithms. *Comput. J.*, 13(3):317–322, 1970.

R. Froeschl, C. Theis, F.P. La Torre, H. Vincke, N . Walter, and A. Sgobba. Radiological Hazard classification of material in CERN's accelerators and the 4 major LHC experiments.

Technical Report CERN-DGS-2012-003-RP-IR v.4. EDMS No: 1184236 v.4, CERN, 2012.

R. Froeschl, M. Magistris, F. Leite Pereira, and C. Theis. Computation of radioactivity in particle accelerators and propagation of uncertainties with the JEREMY code. Technical report, CERN, 2014. CERN-RP–2014–065–REPORTS-TN v.1, EDMS No: 1406467 v.1.

R.O. Gilbert. *Statistical methods for environmental pollution monitoring.* John Wiley and Sons, 1987.

G. Gilmore. *Practical gamma-ray spectrometry.* John Wiley and Sons, 2008.

P. Gy. *Sampling of particulate materials: theory and practice.* Elsevier Scientific Pub. CO., 1979.

P. Gy. *Sampling for analytical purposes.* John Wiley and Sons, 1998.

B. Harding, C. Tremblay, and C. Cousineau. Standard errors: a review and evaluation of standard error estimators using Monte Carlo simulations. *Quant. Meth. Psych.*, 10(2): 107–123, 2014.

T. Hastie, R.J. Tibshirani, and J. Friedman. *The elements of statistical learning. Data mining, inference and prediction.* Springer, 2009.

E. Hecht. *Ondes, Optique et Physique Moderne.* De Boeck, 2007.

IAEA. IAEA-TECDOC-1537. Strategy and methodology for radioactive waste characterization. Technical report, 2007.

IAEA. General Safety Guide No. GSG-1. Classification of radioactive waste. Technical report, 2009a.

IAEA. Nuclear Energy Series No. NW-T-1.18. Determination and use of scaling factors for waste characterization in nuclear power plants. Technical report, 2009b.

IAEA. IAEA Analytical Quality in Nuclear Applications Series No. 48. Determination and interpretation of characteristics limits for radioactive measurements. Decision Threshold, Detection Limit and Limits of the Confidence Interval, 2017.

C.O Ingamells and P. Switzer. A proposed sampling constant for use in geochemical analysis. *Talanta*, 20(6):547–568, 1973.

ISO. ISO 21238. Nuclear energy - Nuclear fuel technology - Scaling factor method to determine the radioactivity of low- and intermediate-level radioactive waste packages generated at nuclear power plants, 2007.

ISO. ISO 11929. Determination of the characteristic limits (decision threshold, detection limit and limits of the confidence interval) for measurements of ionizing radiation — Fundamentals and application, 2010.

ISO. ISO 16966. Nuclear energy - Nuclear fuel technology - Theoretical activation calculation method to evaluate the radioactivity of activated waste generated at nuclear reactors, 2013.

G. James, D. Witten, T. Hastie, and R.J. Tibshirani. *An introduction to statistical learning with applications in R.* Springer, 2017.

JCGM. JCGM 100:2008. Evaluation of measurement data - Guide to the expression of uncertainty in measurement, 2008a.

JCGM. JCGM 101:2008 Evaluation of measurement data - Supplement 1 to the "Guide to the expression of uncertainty in measurement" - Propagation of distributions using Monte Carlo method., 2008b.

JCGM. JCGM 102:2011e Evaluation of measurement data - Supplement 2 to the "Guide to the expression of uncertainty in measurement" - Extension to any number of output quantities., 2011.

V. Kain. Beam dynamics and beam losses - Circular machines. Contribution to the 2014 Joint International Accelerator School: Beam Loss and Accelerator Protection.

J.M. Kirkpatrick, R. Venkataram, and B. Young. Minimum detectable activity, systematic uncertainties, and the ISO 11929 standard. *J. Radioanal. Nucl. Chem.*, 296:1005–1010, 2013.

G.F. Knoll. *Radiation detection and measurements.* John Wiley and Sons, 2010.

K.S. Krane. *Introductory Nuclear Physics.* John Wiley and Sons, Inc., 1987.

D.P. Kroese, T. Taimre, and Z.I. Botev. *Handbook of Monte Carlo methods.* John Wiley and Sons Inc., 2011.

M. Kuhn. Emittance preservation at the LHC, 2013. Master thesis.

F.P. La Torre. *Study of induced radioactivity in proton accelerator facilities.* PhD thesis, University of Berne, 2014.

F.P. La Torre and M. Magistris. ELICA: chemical and physical characterization of TFA cables. Technical Report CERN-RP-2016-087-REPORTS-TN v.1, EDMS No: 1685934 v.1, CERN, 2016.

M. L'Annunziata. *Handbook of radioactivity analysis.* Academic Press, 2012.

C. Lefevre. LHC: The guide, 2009. CERN-Brochure-2009-003-Eng.

M.C. Lépy, A. Pearce, and O. Sima. Uncertainties in gamma-ray spectrometry. *Metrologia*, 52:S123–S145, 2015.

W.G. Madow. On the theory of systematic sampling, III. Comparison of centred and random start systematic sampling. *Ann. Math. Statist.*, 24(1):101–106, 1953.

M. Magistris. *Radiological Characterization of Radioactive Waste at CERN.* PhD thesis, Technischen Universitat Wien, 2008.

M. Magistris and P. Otiougova. Radiological characterization of LEP radiofrequency units for their free-release as conventional waste. Technical Report CERN-RP-2013-060-REPORTS-TN v.1.1, EDMS No: 1299357 v.1.1, CERN, 2013.

M. Magistris and B. Zaffora. Guidelines. Authoritative sampling for the characterization of TFA, metallic waste. Technical Report CERN-RP-2014-113-REPORTS-TN v1.2, EDMS No: 1440750 v.1.2, CERN, 2014.

M. Magistris and B. Zaffora. Optimization of the number of gamma spectrometry measurements for the characterization of waste packages. Technical Report CERN-RP-2015-005-REPORTS-TN v.1, EDMS No: 1470110 v.1, CERN, 2015.

M. Magistris, B. Zaffora, and N. Walter. Radionuclide inventory of metallic, TFA radioactive waste. Technical Report CERN-RP-2015-27-REPORTS-TN v.2.1, EDMS No: 1501107 v.2.1, CERN, 2015.

M. Magistris, R. Froeschl, T. Frosio, and F.P. La Torre. Radionuclide inventory of TFA cables. Technical Report CERN-RP-2016-114-REPORTS-TN v.3, EDMS No: 1703324 v.3, CERN, 2017a.

M. Magistris, C. Theis, and Y. Pira. Radionuclide inventory of TFA magnets. Technical Report CERN-RP-2016-112-REPORTS-TN v.3, EDMS No: 1700907 v.3, CERN, 2017b.

P. Mazzoldi, M. Nigro, and C. Voci. *Fisica 2 - Elettromagnetismo, Onde e Ottica*. Edises, 1998.

METAS. Cours en métrologie. Cours de base en incertitude de mesure, 2012.

R. Michaud and B. Zaffora. Dossier de demande d'acceptation du lot de dechets historiques de blocs metalliques compactes. Technical Report CERN-RP-2013-082-REPORTS-TN v.2, EDMS No: 1310984 v.2, CERN, 2013.

A. Mitaroff and M. Silari. The CERN-EU high-energy reference field (CERF) facility for dosimetry at commercial flight altitudes and in space. *Rad. Prot. Dos.*, 102(1):7–22, 2002.

J.C. Myers. *Geostatistical error management: quantifying uncertainty for environmental sampling and mapping*. John Wiley and Sons, 1997.

A. Nagy and B. Zaffora. Optimization of γ-spectrometry analysis for the characterization of TFA waste packages. Technical Report CERN-RP-2016-028-REPORTS-TN v.1, EDMS No: 1584185 v.1, CERN, 2016.

J.A. Nelder and R. Mead. A simplex method for function minimization. *Comput. J.*, 7(4): 308–313, 1965.

D.J. Olive. Prediction intervals for regression models. *Comput. Stat. Data Anal.*, 51(6): 3115–3122, 2007.

P.N. Ostroumov, V.N. Aseev, and B. Mustapha. Beam loss studies in high-intensity heavy-ion linacs. *Phys. Rev. ST Accel. Beams*.

P. Patnaik. *Dean's analytical chemistry handbook*. McGraw-Hill, 2004.

F.F. Pitard. *Pierre Gy's sampling theory and sampling practice. Heterogeneity, sampling correctness, and statistical process control*. CRC Press, 1993.

M.A. Plum. Beam loss mechanisms in high intensity linacs. In *HB2012: Proceedings of the 52nd ICFA Advanced Beam Dynamics Workshop on High-Intensity and High-Brightness Hadron Beams*, 2012.

R CORE TEAM. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, 2014. URL `http://www.R-project.org/`.

J. Ranft. Dual parton model at cosmic ray energies. *Phys. Rev. D Part. Fields*, 51(1): 64–84, 1995.

H. Ross, J.E. Noakes, and J.D. Spaulding. *Liquid scintillation counting and organic scintillators*. Lewis Publisher Inc., 1991.

G. Van Rossum and F.L. Drake. *Python reference manual*. PythonLabs, 2001. URL `https://www.python.org/`.

M.E. Schlesinger, M. King, K. Sole, and W. Davenport. *Extractive metallurgy of copper*. Elsevier, 2011.

C. Theis and H. Vincke. Actiwiz - Optimizing your nuclide inventory at proton accelerators with a computer code. In *International Conference on Radiation Shielding (ICRS-12)*, 2012.

P. Verrier. Plastifiants. *Techniques de l'ingénieur. Adjuvants des plastiques*, Base documentaire: TIB138DUO.(A3231 V1), 1992.

H. Vincke and C. Theis. Radiological characterization of cable materials used in high-energy accelerators at CERN. Technical Report CERN-RP-2015-081-REPORTS-TN v.2, EDMS No: 1541541 v.2, CERN, 2015.

H. Wiedemann. *Particle accelerator physics*. Springer-Verlag, 2007.

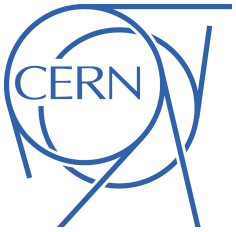
E. Wilson. *An introduction to particle accelerators*. Clarendon Press, 2001.

B. Zaffora. Radiological characterization of printed circuit boards for future elimination, 2014. Master thesis.

B. Zaffora, M. Magistris, G. Saporta, and F.P. La Torre. Statistical sampling applied to the radiological characterization of historical waste. *EPJ Nuclear Sci. Technol.*, 2(34), 2016.

B. Zaffora, M. Magistris, J.-P. Chevalier, C. Luccioni, G. Saporta, and L. Ulrici. A new approach to characterize very-low-level radioactive waste produced at hadron accelerators. *Appl. Rad. Isot.*, 122:141–147, 2017a.

B. Zaffora, M. Magistris, G. Saporta, and J.-P. Chevalier. Uncertainty quantification applied to the radiological characterization of radioactive waste. *Appl. Rad. Isot.*, 127: 142–149, 2017b.



Biagio ZAFFORA

Statistical analysis for the radiological characterization of radioactive waste in particle accelerators

Abstract :

This thesis introduces a new method to characterize metallic very-low-level radioactive waste produced at the European Organization for Nuclear Research (CERN). The method is based on: 1. the calculation of a preliminary radionuclide inventory, which is the list of the radionuclides that can be produced when particles interact with a surrounding medium, 2. the direct measurement of γ emitters and, 3. the quantification of pure-α, pure-β and low-energy X-ray emitters, called difficult-to-measure (DTM) radionuclides, using the so-called scaling factor (SF), correlation factor (CF) and mean activity (MA) techniques. The first stage of the characterization process is the calculation of the radionuclide inventory via either analytical or Monte Carlo codes. Once the preliminary radionuclide inventory is obtained, the γ-emitting radionuclides are measured via γ-ray spectrometry on each package of the waste population. The major γ-emitter, called key nuclide (KN), is also identified. The scaling factor method estimates the activity of DTM radionuclides by checking for a consistent and repeated relationship between the key nuclide and the activity of the difficult to measure radionuclides from samples. If a correlation exists the activity of DTM radiodionuclides can be evaluated using the scaling factor otherwise the mean activity from the samples collected is applied to the entire waste population. Finally, the correlation factor is used when the activity of pure-α, pure-β and low-energy X-ray emitters is so low that cannot be quantified using experimental values. In this case a theoretical correlation factor (CF) is obtained from the calculations to link the activity of the radionuclides we want to quantify and the activity of the key nuclide. The thesis describes in detail the characterization method, shows a complete case study and describes the industrial-scale application of the characterization method on over 1000 m^3 of radioactive waste, which was carried out at CERN between 2015 and 2017.

Keywords : Particle accelerator, radioactive waste, statistical analysis, scaling factor, activation.

Résumé :

Ce travail de thèse introduit une nouvelle méthode pour la caractérisation radiologique des déchets métalliques très faiblement radioactifs produits au sein de l'Organisation Européenne pour la Recherche Nucléaire (CERN). La méthode se base sur: 1. le calcul des radionucléides en présence, i.e. les radionucléides qui peuvent être produits lors de l'interaction des particules avec la matière et avec les structures environnantes les accélérateurs, 2. la mesure directe des émetteurs γ et, 3. la quantification des émetteurs α et β purs et de rayons X de faible énergie, appelés radionucléides difficile-a-mesurer (DTM), en utilisant les méthodes dites des "scaling factor" (SF), "correlation factor" (CF) et activité moyenne (MA). La première phase du processus de caractérisation est le calcul des radionucléides en présence à l'aide de codes de calcul analytiques ou Monte Carlo. Après le calcul de l'inventaire radiologique, les radionucléides émetteurs γ sont mesurés par spectrométrie γ dans chaque colis de la population. L'émetteur γ dominant, appelé "key nuclide" (KN) est identifié. La méthode dite des "scaling factors" permet d'estimer l'activité des radionucléides DTM après évaluation de la corrélation entre l'activité des DTM et l'activité de l'émetteur γ dominant obtenue à partir d'échantillons. Si une corrélation existe, l'activité des radionucléides DTM peut être évaluée grâce à des facteurs de corrélation expérimentaux appelés "scaling factors", sinon l'activité moyenne obtenue à partir d'échantillons prélevés dans la population est attribuée à chaque colis. Lorsque les activités des émetteurs α et β purs et des émetteurs X de faible énergie ne peuvent pas être estimées par mesure analytique la méthode des "correlation factors" s'applique. La méthode des "correlation factors" se base sur le calcul de corrélations théoriques entre l'émetteur γ dominant et les radionucléides de très faible activité. Cette thèse décrit en détail la nouvelle technique de caractérisation radiologique, montre un cas d'application complet et présente les résultats de l'industrialisation de la méthode ayant permis la caractérisation radiologique de plus de 1000 m^3 de déchets radioactifs au CERN entre 2015 et 2017.

Keywords : Accélérateur de particules, déchets radioactifs, analyse statistique, scaling factor, activation.