TKK Dissertations 129
Espoo 2008

ANALYSIS METHODS FOR AIRBORNE RADIOACTIVITY

Doctoral Dissertation

Jarmo J. Ala-Heikkilä

Helsinki University of Technology
Faculty of Information and Natural Sciences
Department of Engineering Physics

TKK Dissertations 129
Espoo 2008

ANALYSIS METHODS FOR AIRBORNE RADIOACTIVITY

Doctoral Dissertation

Jarmo J. Ala-Heikkilä

Dissertation for the degree of Doctor of Science in Technology to be presented with due permission of the Faculty of Information and Natural Sciences for public examination and debate in Auditorium K216 at Helsinki University of Technology (Espoo, Finland) on the 20th of August, 2008, at 12 noon.

Helsinki University of Technology
Faculty of Information and Natural Sciences
Department of Engineering Physics

Teknillinen korkeakoulu
Informaatio- ja luonnontieteiden tiedekunta
Teknillisen fysiikan laitos

Distribution:
Helsinki University of Technology
Faculty of Information and Natural Sciences
Department of Engineering Physics
P.O. Box 4100
FI - 02015 TKK
FINLAND
URL: http://www.tkk.fi/Units/AES/
Tel. +358-9-451 3198
Fax +358-9-451 3195
E-mail: jarmo.ala-heikkila@tkk.fi

ISBN 978-951-22-9439-8
ISBN 978-951-22-9440-4 (PDF)
ISSN 1795-2239
ISSN 1795-4584 (PDF)
URL: http://lib.tkk.fi/Diss/2008/isbn9789512294404/

TKK-DISS-2491

Yliopistopaino
Helsinki 2008

ABSTRACT OF DOCTORAL DISSERTATION

HELSINKI UNIVERSITY OF TECHNOLOGY
P. O. BOX 1000, FI-02015 TKK
http://www.tkk.fi

Author Jarmo Johannes Ala-Heikkilä

Name of the dissertation
Analysis Methods for Airborne Radioactivity

Manuscript submitted 11.02.2008
Manuscript revised 23.05.2008
Date of the defence 20.08.2008

☒ Monograph ☐ Article dissertation (summary + original articles)

Faculty	Faculty of Information and Natural Sciences
Department	Department of Engineering Physics
Field of research	Radiation physics
Opponent(s)	Prof. Stanley G. Prussin
Supervisor	Prof. Rainer Salomaa
Instructor	D.Sc. Pertti Aarnio

Abstract

High-resolution gamma-ray spectrometry is an analysis method well suitable for monitoring airborne radioactivity. Many of the natural radionuclides and a majority of anthropogenic nuclides are prominent gamma-ray emitters. With gamma-ray spectrometry different radionuclides are readily observed at minute concentrations that are far from health hazards. The gamma-ray spectrometric analyses applied in air monitoring programmes can be divided into particulate measurements and gas measurements. In this work, methods applicable for particulate sample analysis have been presented, implemented in analysis software, and evaluated with a wide variety of cases. Our goal has been to develop a collection of tools that enables a complete quantitative explanation of all components of a measured gamma-ray spectrum with a minimum of user intervention.

In a high-resolution spectrum, all essential information is contained in the peaks. Most of them are full-energy peaks that can be explained with gamma-ray lines in a reference library. Among the full-energy peaks, however, there are annihilation escape peaks, X-ray escape peaks, coincidence sum peaks, and random sum peaks that are not tabulated. Calculation methods for these special peaks are presented, implemented in SHAMAN, and evaluated in this work.

Expert system SHAMAN combines a comprehensive ENSDF-based reference library with an inference engine that applies pruning rules to select the acceptable candidate nuclides, utilizing a collection of calculational methods. Its performance with air filter spectra is evaluated in this work with four different spectrum sets. The performance of SHAMAN is also benchmarked against another identification software in completely automated analysis.

In summary, expert system SHAMAN is well applicable at any organization where environmental radioactivity is monitored. The system can also be utilized in other applications, but the current rule base has been tailored most comprehensively for air filter spectra. The analysis capabilities of the system can be improved in other applications of gamma-ray spectrometry by tuning the analysis parameters of the system. Naturally, there is still room for improvement in the methodology of SHAMAN, but already with the currently implemented features it is a reliable expert system for nuclide identification.

Keywords gamma-ray spectrometry, airborne radioactivity, radionuclide identification
ISBN (printed) 978-951-22-9439-8
ISSN (printed) 1795-2239
ISBN (pdf) 978-951-22-9440-4
ISSN (pdf) 1795-4584
Language English
Number of pages 188 p.
Publisher TKK
Print distribution TKK

VÄITÖSKIRJAN TIIVISTELMÄ	TEKNILLINEN KORKEAKOULU PL 1000, 02015 TKK http://www.tkk.fi

Tekijä Jarmo Johannes Ala-Heikkilä

Väitöskirjan nimi

Analyysimenetelmiä ilman radioaktiivisuuden valvontaan

Käsikirjoituksen päivämäärä 11.02.2008	Korjatun käsikirjoituksen päivämäärä 23.05.2008

Väitöstilaisuuden ajankohta 20.08.2008

☒ Monografia	☐ Yhdistelmäväitöskirja (yhteenveto + erillisartikkelit)

Tiedekunta	Informaatio- ja luonnontieteiden tiedekunta
Laitos	Teknillisen fysiikan laitos
Tutkimusala	Säteilyfysiikka
Vastaväittäjä(t)	Prof. Stanley G. Prussin
Työn valvoja	Prof. Rainer Salomaa
Työn ohjaaja	TkT Pertti Aarnio

Tiivistelmä

Korkean resoluution gammaspektrometria on analyysimenetelmä, joka soveltuu hyvin ilman radioaktiivisuuden valvontaan, sillä useimmat luonnolliset ja ihmisen tekemät radionuklidit lähettävät gammasäteilyä. Käytännössä gammaspektrometrialla on mahdollista havaita minimaalisia aktiivisuuspitoisuuksia, jotka eivät ole lähelläkään terveysriskiä. Gammaspektrometriset analyysit voidaan jaotella hiukkas- ja kaasunäytemittauksiin. Tässä työssä esitellään menetelmiä hiukkasnäytteiden analyysiin, ohjelmoidaan ne tietokoneohjelmaksi ja arvioidaan niiden käyttökelpoisuus suuren testispektrijoukon avulla. Tavoitteena on kehittää sellainen työkalukokoelma, jonka avulla kaikki gammaspektrin komponentit voidaan selittää kvantitatiivisesti ja joka toimii mahdollisimman automaattisesti.

Korkean resoluution gammaspektrissä kaikki tieto on spektrin piikeissä. Suurin osa piikeistä on täysenergiapiikkejä, joita vastaavat gammaviivat löytyvät referenssikirjastosta. Täysenergiapiikkien seassa on kuitenkin annihilaatiopakopiikkejä, röntgenpakopiikkejä, koinsidenssisummapiikkejä ja satunnaissummapiikkejä, joita ei ole taulukoitu. Tässä työssä esitellään, ohjelmoidaan ja evaluoidaan laskentamenetelmiä näille erikoispiikeille.

Työssä on kehitetty asiantuntijajärjestelmä Shamania. Se yhdistää kattavan ENSDF-pohjaisen referenssikirjaston ja laskentamenetelmät päättelykoneeseen, joka käyttää karsintasääntöjä analysoitavan näytteen oikean nuklidikoostumuksen päättelemiseen. Shamanin suorituskyky ilmafiltterien analyysissa evaluoidaan tässä työssä neljän laajan spektrikokoelman avulla. Lisäksi suorituskykyä arvioidaan vertailemalla tuloksia toisen tunnistusohjelmiston saavuttamiin.

Yhteenvetona voidaan todeta, että Shaman soveltuu mainiosti luonnon radioaktiivisuuden valvontatehtävään. Sitä voidaan käyttää myös muissa sovelluskohteissa, joissa pyritään täydelliseen gammaspektrin tulkintaan, mutta nykyversion tietämyskanta on parhaiten räätälöity ilmafiltterispektreille. Analyysikykyä muissa sovelluksissa voidaan parantaa ohjelmiston parametreja säätämällä. Myös Shamanin metodologiassa on kehittämisen varaa, mutta jo nykyisillä laskentamenetelmillä se on luotettava asiantuntijajärjestelmä radionuklidien tunnistukseen.

Asiasanat gammasäteily, gammaspektrometria, radioaktiivisuusvalvonta, radionuklidien tunnistus

ISBN (painettu) 978-951-22-9439-8	ISSN (painettu) 1795-2239
ISBN (pdf) 978-951-22-9440-4	ISSN (pdf) 1795-4584
Kieli englanti	Sivumäärä 188 s.

Julkaisija TKK

Painetun väitöskirjan jakelu TKK

☒ Luettavissa verkossa osoitteessa http://lib.tkk.fi/Diss/2008/isbn9789512294404/

Abstract

High-resolution gamma-ray spectrometry is an analysis method well suitable for monitoring airborne radioactivity. Many of the natural radionuclides as well as a majority of anthropogenic nuclides are prominent gamma-ray emitters. With gamma-ray spectrometry different radionuclides are readily observed at minute concentrations that are far from health hazards. Thus, measurements during normal conditions and minor anthropogenic releases can be utilized to increase sensitivity of analysis methods and to learn to distinguish between source terms from different release types.

The gamma-ray spectrometric analyses applied in air monitoring programmes can be divided into particulate measurements and gas measurements, of which the former category is technologically more mature. In this work, methods applicable for particulate sample analysis have been presented, implemented in analysis software, and evaluated with a wide variety of cases. Our goal has been to develop a collection of tools that enables a complete quantitative explanation of all components of a measured gamma-ray spectrum with a minimum of user intervention.

The importance of accurate and precise calibrations cannot be overestimated. Energy and peak efficiency calibrations are fundamental for spectrum analysis, since they are needed to convert the peak positions and areas to energies and emission rates and further to nuclide identifications and quantifications. Use of a shape calibration can also be recommended for its stabilizing effect on peak search and area determination. If measurements are made in a close geometry, a total efficiency calibration is also needed in order to enable corrections for true coincidence summing effects.

In a high-resolution spectrum, all essential information is contained in the peaks. Most of them are full-energy peaks that can be explained with gamma-ray lines in a reference library. Among the full-energy peaks, however, there are annihilation escape peaks, X-ray escape peaks, coincidence sum peaks, and random sum peaks that are not tabulated. They need to be explained quantitatively if a complete interpretation of a measured spectrum is aimed at. Calculation methods for these special peaks are presented, implemented, and evaluated in this work.

Expert system Shaman combines a comprehensive ENSDF-based reference library with an inference engine that applies pruning rules to select the acceptable candidate nuclides, and with a collection of calculational methods including the above-mentioned ones. Its performance with particulate air filter spectra has been evaluated in this work with four different spectrum sets. Two of the sets have been produced at different times by a global monitoring network that has been set up for the Comprehensive Nuclear-Test-Ban Treaty (CTBT), and the two other sets, measured after two minor anthropogenic releases, by the Finnish monitoring network run by STUK (Säteilyturvakeskus, Finnish Radiation and Nuclear Safety Authority).

In the first CTBT-related set of 250 spectra the peak analysis was made with tailor-made analysis software based on Genie. Its output was analyzed with two SHAMAN versions and they both obtained a nuclide identification percentage and a peak identification percentage of approximately 99 % each. The spurious nuclide percentage was 4–7 %, translating to 0.4–0.7 spurious identifications per spectrum. These results were obtained in a completely automated analysis mode, so they can be considered excellent.

The second CTBT-related set included 6 161 spectra that were not analyzed manually unlike the first spectrum set, only with an automated UNISAMPO–SHAMAN pipeline system. Without a manual analysis we presented uncorrected peak identification percentages and uncorrected average numbers of spurious nuclides of 96.4 % and 0.96, respectively. Both figures are less ideal than for the previous spectrum set, but the explanation is that a smaller peak search threshold has been utilized in the peak analysis phase. Therefore, the average number of peaks is much larger than that in the first set and especially the share of small peaks has increased. The performance level can still be considered very good, even if there is some room for improvement.

The spectrum sets of the national network had been measured after minor releases from Novaya Zemlya in August 1987 and Sosnovyy Bor in March 1992. Both spectrum sets were analyzed with SAMPO in interactive mode and then identified with SHAMAN, so the results are not directly comparable to the automated analysis results referred to above. However, these two spectrum sets showed that the identification performance of SHAMAN is not degraded in cases where air filter spectra contain fission and activation products in addition to natural radioactivity.

The performance of SHAMAN was benchmarked against another identification software based on Genie in two intercomparisons. In both cases, the results from automated analysis were compared, thus neglecting the effect of a human analyst that can compensate for errors in automated results. The first intercomparison truly measured the difference in nuclide identification, since the same peak results were fed into both identification implementations. In the second intercomparison two automated analysis pipelines were compared where also the peak analysis results are different due to different software. In both intercomparisons, SHAMAN had a clear advantage in nuclide identification and peak explanation. The statistically inevitable disadvantage, a larger number of spurious nuclides, remained under control.

In summary, the UNISAMPO–SHAMAN pipeline system is well applicable at any organization where environmental radioactivity is monitored. The system can also be utilized in other applications, but the current rule base of expert system SHAMAN has been tailored most comprehensively for air filter spectra. The analysis capabilities of the system can be improved in other applications of gamma-ray spectrometry by tuning the analysis parameters of the system. This should not require but a modest effort of an end-user who has the application-specific expertise and has some scripting experience since a major part of SHAMAN's "expertise" or "intelligence" is in its operating scripts and macros. Naturally, there is still room for improvement in the methodology of SHAMAN, but with the currently implemented features it is a reliable expert system for nuclide identification.

Preface

This thesis presents the results from a long-lasting development project where the methodology of expert system SHAMAN has been improved and its performance has been evaluated at various stages of development. The project has been financed primarily by Helsinki University of Technology (TKK), but over the years financial support has also been obtained from the Ministry of Trade and Industry and the Ministry for Foreign Affairs of Finland, as well as from the US Department of Defense through the visiting scientist programme of the Prototype International Data Centre (PIDC), Arlington, Virginia.

The author is most grateful for this long-term financial support. Especially the extreme patience of my supervisor Prof. R. Salomaa is admirable. I also wish to state my gratitude for the personal grants from the Jenny and Antti Wihuri Foundation and the Swedish Academy of Technical Sciences in Finland (Wilhelm Guerillot Fund) in mid-1990's.

The assistance and co-operation of the radionuclide staff at the PIDC during my two visits in 1996–98 is highly appreciated. The following persons contributed to the success in one way or another: Dr. N. Duncan, Dr. R. Mason, Dr. W. Evans, Ms. J. Bohlin, Mr. G. Novosel, Mr. E. Pratt, Mr. J. Bohner, Mr. A. Bennett, Mr. J. Nicholson, Dr. K. Biegalski, Dr. S. Biegalski, Dr. L. Vladimirski, Mr. M. Pickering, Mr. M. Wasim, and Mr. D. Williams.

Over the years, we have enjoyed a fruitful collaboration with the Security Technology Laboratory (formerly Aerosol Laboratory) of STUK headed by Dr. H. Toivonen, as well as the Finnish National Data Centre (FiNDC) hosted by STUK and headed consecutively by Mr. M. Nikkinen, Mr. A. Isolankila, and Mr. M. Moring.

A major part of the research has been done in the Radiation Physics Group of the Advanced Energy Systems laboratory at TKK. The group is not big but its expertise on gamma-ray spectrometry is state-of-the-art. The group leader and my instructor Dr. P. Aarnio has introduced an extremely interesting research subject to me and has led the work with scientific rigor through the years. Mr. T. Hakulinen is the father of SHAMAN who has built and continues to build an exemplarily robust computer code, a solid foundation that allows for further development. Mr. M. Nikkinen has contributed to the SHAMAN project in his Master's thesis and later by developing the spectrum analysis software UNISAMPO, the preferred front-end for SHAMAN. Prof. Emer. J. Routti, the founder of this group, has shown continued interest in our work, inspiring us to carry on. Additionally, a number of smart summer students have contributed to the research of our group at various stages.

The working atmosphere in the Advanced Energy Systems laboratory has been most pleasant and inspiring, but nevertheless, I would like to emphasize the contribution of my family. Without its counterbalancing effect this work may have been completed earlier, but my life would have been much less rewarding. Therefore, I wish to dedicate this work to my dear wife Marika and our four wonderful children, Ida, Fanny, Anton, and Sofia.

Jarmo Ala-Heikkilä
Espoo, May 23, 2008

List of Publications and Author's Contribution

This thesis is a monograph but it is based on previous publications of the author that are listed in the References beginning from p. 163. Roman uppercase numbering has been used for these publications in references and an Arabic one for publications by other authors. The most essential publications are the following:

[IV] J.J. Ala-Heikkilä, *Expert System for Nuclide Identification in Environmental Gamma Spectra.* Report TKK-F-B159, Otaniemi 1995.

[VIII] J.J. Ala-Heikkilä, T.T. Hakulinen, P.A. Aarnio, M.T. Nikkinen, H.I.K. Toivonen, *Evaluation of Expert System* SHAMAN *in Processing Gamma-Ray Spectra at the Comprehensive Test Ban Treaty Prototype International Data Center.* Report TKK-F-B171, Otaniemi 1997.

[IX] P.A. Aarnio, J.J. Ala-Heikkilä, T.T. Hakulinen, M.T. Nikkinen, "Application of the Nuclide Identification System SHAMAN in Monitoring the Comprehensive Test Ban Treaty". Journal of Radioanalytical and Nuclear Chemistry **235** 1–2 (1998) 95–103.

[XI] P.A. Aarnio, J.J. Ala-Heikkilä, T.T. Hakulinen, M.T. Nikkinen, "The Nuclide Identification System SHAMAN in the Verification of the Comprehensive Nuclear-Test-Ban Treaty". Journal of Radioanalytical and Nuclear Chemistry **248** 3 (2001) 587–593.

[XIV] P.A. Aarnio, J.J. Ala-Heikkilä, T.K. Ansaranta, T.T. Hakulinen, "Analysis Pipeline for Air Filter Gamma-Ray Spectra from the CTBT Verification Network". Journal of Radioanalytical and Nuclear Chemistry **263** 1 (2005) 251–257.

[XVII] P.A. Aarnio, J.J. Ala-Heikkilä, T.T. Hakulinen, "Performance of UNISAMPO–SHAMAN with Gamma-Ray Spectra Containing Known Traces of Fission Products". Journal of Radioanalytical and Nuclear Chemistry **276** 2 (2008) 455–460.

The author of this work has been the principal author of these publications and a co-author in the other Roman-numbered publications in the References. Otherwise, the author has contributed by developing various analysis methods for expert system SHAMAN and verifying them. Additionally, the author has had the main responsibility for performance evaluation of SHAMAN.

Abbreviations and Acronyms

ADC	Analog to Digital Converter (in a multichannel analyzer)
ANSI	American National Standards Institute
ARR	Automated Radionuclide Report (produced by the CTBTO)
BEGe	Broad-Energy Germanium (commercial detector brand)
BGO	Bismuth Germanate, $Bi_4Ge_3O_{12}$ (detector material)
BNL	Brookhaven National Laboratory
CERN	Organisation Européenne pour la Recherche Nucléaire (European Organization for Nuclear Research)
CMR	Center for Monitoring Research
CPU	Central Processing Unit (in a computer)
CTBT	Comprehensive Nuclear-Test-Ban Treaty
CTBTO	Comprehensive Nuclear-Test-Ban Treaty Organization
CZT	Cadmium Zinc Telluride (detector material)
DE	Double Escape (gamma-ray spectrum peak)
DIN	Deutsche Institut für Normung (German Standards Authority)
EC	Electron Capture (decay mode)
ENSDF	Evaluated Nuclear Structure Data File
EOI	End of Irradiation
FWHM	Full Width at Half Maximum
FiNDC	Finnish National Data Centre
HPGe	High Purity Germanium (detector material)
IAEA	International Atomic Energy Agency
IDC	International Data Centre (part of CTBTO)
IMS	International Monitoring System (part of CTBTO)
IT	Internal Transition (decay mode)
LLD	Lower Level Discriminator
MC	Monte Carlo simulation
MCA	Multichannel Analyzer
MDA	Minimum Detectable Activity
MDC	Minimum Detectable Concentration
NaI	Sodium Iodide (detector material)
NDC	National Data Centre (related with CTBTO)
NNDC	National Nuclear Data Center (at BNL)
NPP	Nuclear Power Plant
NRC	Nuclear Regulatory Commission
NUDAT	Nuclear structure and decay Data
OECD	Organization for Economic Cooperation and Development
PIDC	Prototype International Data Centre
QC	Quality Control
RGL	Report Generating Language
ROI	Region of Interest (in spectrum analysis)
SCA	Single Channel Analyzer
SE	Single Escape (gamma-ray spectrum peak)
SQL	Structured Query Language
STUK	Säteilyturvakeskus (Finnish Radiation and Nuclear Safety Authority)
ToRI	Table of Radioactive Isotopes

Contents

Chapter 1

Introduction

Environmental radioactivity surveillance networks traditionally utilize count rate meters located in a network around a nuclear facility or around a territory. This kind of a surveillance network can easily be automated because the output from the detectors is a single parameter, count rate, as a function of time [20, p. 448]. Count rate monitoring implements the first level of radiation protection of the public.

A network of count rate meters is complemented with sampling of air, water, soil and food with consequent analysis using various radioanalytical methods. Airborne radioactivity, the subject of this work, is sampled at air filtering stations. At these stations, air is pulled through a filter typically for one or several days after which the filter is measured with a gamma-ray spectrometer. This kind of a surveillance network is very sensitive for detecting airborne radioactivity, because in a high-resolution gamma-ray spectrum of a filter sample, peaks of anthropogenic radionuclides are readily distinguished from those of natural background nuclides. It is possible to determine the source of anthropogenic nuclides when they are reliably identified from an air filter spectrum. Using meteorological transport models and calculated activity ratios, it is possible to verify the source and assess the release time.

Air filter sampling has been an application that has required human labor to run the stations and considerable expertise to analyze the gamma-ray spectra. However, more automated sampling stations have become available recently. A network of automated or semi-automated sampling stations may be complemented with an automated analysis pipeline. This can be implemented in a single workstation that runs suitable software packages for receiving and storing the gamma-ray spectra measured at the stations, for analyzing the spectra, and for generating alarms if anything out of the ordinary is observed.

As an international endeavor, a network of airborne radioactivity measurement stations is being built in order to monitor compliance with the Comprehensive Nuclear-Test-Ban Treaty (CTBT) [21]. This network is utilized in parallel with networks of seismic, hydroacoustic and infrasound stations. Since there will be 80 radionuclide stations around the globe, many of them in distant locations, automation of the stations has been an important design goal. The same goal applies to analysis of measurement data at the CTBT Organization (CTBTO) due to large volumes of data and requirements for objectivity. The CTBT-related radionuclide monitoring issues are elaborated in Ch. 2.

When monitoring airborne radioactivity with a spectrometer system, different radionuclides are detected in concentrations that are orders of magnitude below hazardous levels. However, even the smallest observation may be useful for verifying whether a nuclear facility obeys regulations or whether someone has performed a clandestine nuclear test. The aim is to distinguish between different types of anthropogenic releases on an inevitable background of natural radionuclides. These releases can be (a) normal airborne effluents from operating nuclear facilities, or (b) accident-level releases from these facilities, or (c) releases from nuclear weapon tests. For the CTBTO, the focus is on the last category of releases, but radiation protection organizations would be interested in any releases. However, radionuclide contents of the atmosphere in normal conditions, and the daily or seasonal variations of the contents, must be known in order to be able to detect the anomalies caused by releases. These issues are elaborated in Ch. 3.

Different methods for high-resolution gamma-ray spectrometry and radionuclide identification have been developed since the early 1960's when germanium detectors became commercially available [22]. In principle, the problem is easy to formulate: find the peaks from a measured spectrum and explain them with a reference library. However, complications arise in practice if a complete explanation of a gamma-ray spectrum is the aim. In spectrum analysis, the basic challenge is to find all true peaks above a noisy baseline, even the smallest ones, without an excessive number of false peaks. In radionuclide identification based on the peaks found, the basic challenge is to infer the nuclide composition that explains the peaks sufficiently well, when there are over 3 600 nuclides and 80 000 gamma-ray and X-ray lines in a comprehensive reference library [23] and some of the spectrum peaks are not tabulated at all, e.g., escape and sum peaks.

This work is concentrated on gamma-ray spectrometric methods implemented in SHAMAN, an expert system for radionuclide identification developed at the Helsinki University of Technology (TKK) since the late 1980's [III, 1–5]. The design goal of SHAMAN has been to provide the correct explanation for all spectrum peaks utilizing a comprehensive reference library, thus being directly suitable for various applications of gamma-ray spectrometry. Analysis methods of SHAMAN are elaborated in Ch. 4 and the current implementation of the expert system is presented in Ch. 5.

A major part of this work has been devoted to performance evaluation of SHAMAN. Two basic methods, baseline analysis and benchmarking, have been utilized in this process. The former essentially means comparing results from a new version to those of an older version, the baseline, and it is the basic tool in everyday development work. The latter means comparing obtained results to those from another source, usually careful manual analysis or an alternative software package, and this method is useful after major revisions and applicable for publication purposes. SHAMAN's main application has been analysis of air filter spectra which is reflected in the case studies of Ch. 6. In both baseline analysis and benchmarking, air filter spectra from the national surveillance network of Finland as well as from the global CTBT monitoring network have been analyzed with SHAMAN. Direct performance comparisons with other identification systems are presented in Ch. 7.

Chapter 2

Comprehensive Nuclear-Test-Ban Treaty — CTBT

After decades of diplomatic negotiations and a series of more limited treaties, the Comprehensive Nuclear-Test-Ban Treaty (CTBT) [21] was opened for signature on September 24, 1996. An organization, the Preparatory Commission for the Establishment of the CTBT Organization, was established in 1997 to operate as a predecessor of the actual CTBTO that will monitor compliance with the treaty when it has entered into force [24]. The treaty specifies the prerequisites for its entry into force, most importantly lists 44 states that possess nuclear technology and are therefore requested to ratify the treaty. For brevity, the organization is called the CTBTO in the following.

Despite the proliferation risks of nuclear weapons and health consequences of nuclear tests were known since the 1950's, the major obstacle for a comprehensive ban was the lack of credible methods to verify compliance with such an agreement. The Cold War started to defrost in the late 1980's with political changes in the Soviet Union and ended with the collapse of the former superpower in 1991. This made room for continued negotiations on the test ban issue and resulted in the CTBT in 1996.

In this Chapter, the verification methods of the CTBT are presented briefly. The main emphasis is on airborne radionuclide monitoring that is highlighted throughout this work. The other technologies and the political history of the treaty, as interesting as they may be, are left to experts in those areas.

2.1 Verification of the CTBT

The CTBTO has its headquarters in Vienna, Austria. It operates the International Data Centre (IDC) where data from worldwide networks, measuring seismic, hydroacoustic, infrasound, and radionuclide signals, are analyzed. These networks are known as the International Monitoring System (IMS) and they are an integral part of the CTBT, enabling the State Signatories to verify each other's compliance with the treaty. A map of the IMS monitoring facilities is shown in Fig. 2.1.

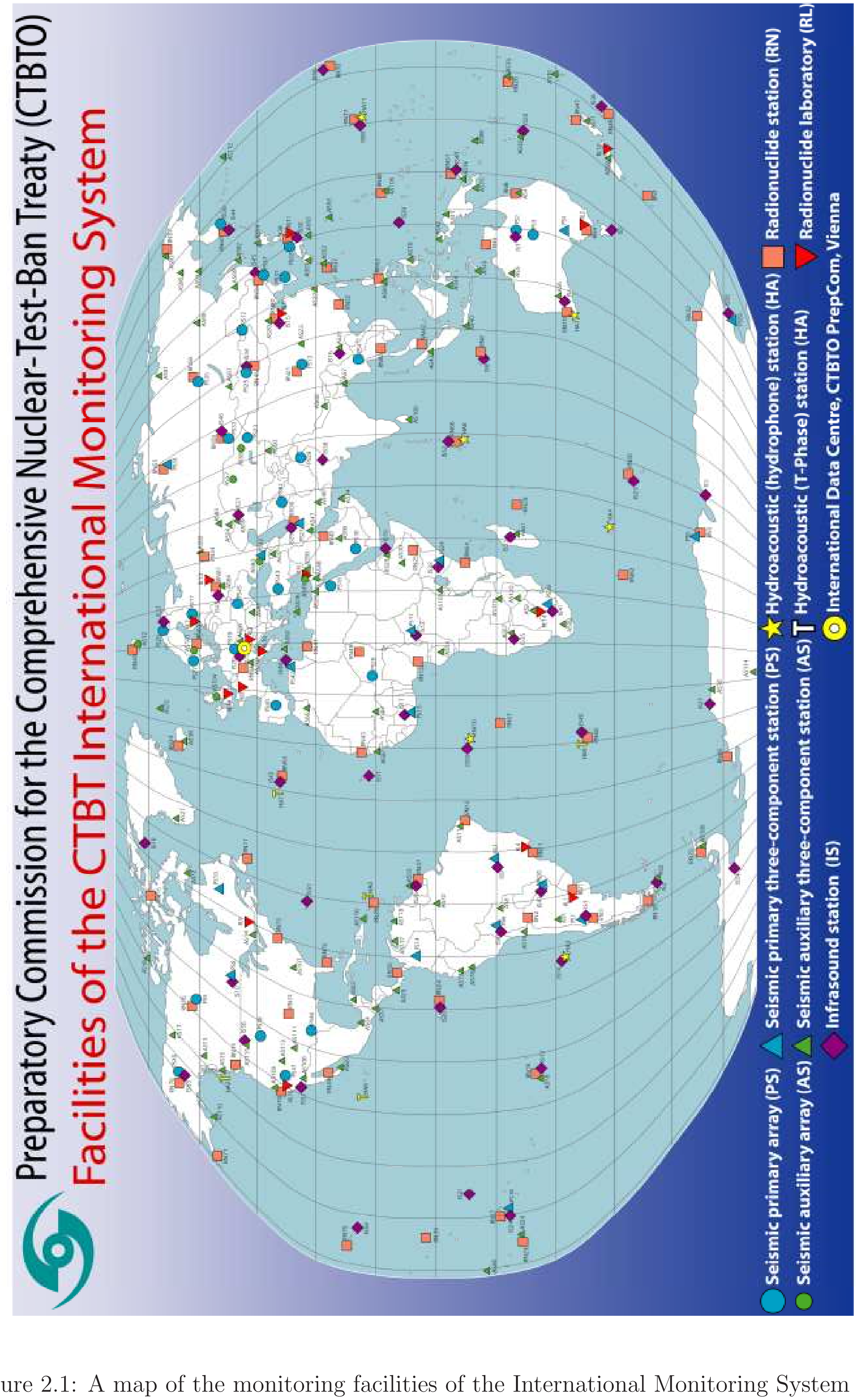

Figure 2.1: A map of the monitoring facilities of the International Monitoring System [24].

It is obvious that the monitoring facilities have been located with a maximum coverage of the globe as the aim. The number of facilities and their locations were specified already during CTBT negotiations. In support of the diplomatic negotiations, the scientific community developed the following four-technology concept where the different networks complement each other:

- The *seismic network* detects periodic vibrations of Earth's crust, following underground and seafloor detonations.
- The *hydroacoustic network* detects acoustic waves in the oceans, following underwater detonations.
- The *infrasound network* detects very low-frequency sound waves in the atmosphere, following atmospheric and near-surface detonations.
- The *radionuclide network* detects radioactive particles, following atmospheric detonations and also underground and underwater tests; half of the stations also detect noble gases that are nearly impossible to contain.

All technologies have their limitations in distinguishing between natural and man-made events. For example, the radionuclide network daily observes natural airborne radionuclides like ^{7}Be and the gamma-emitting radon daughters, sometimes also releases from nuclear power plants and accelerators, so the challenge is to observe CTBT-relevant signals above this background. However, when utilizing the four different technologies in parallel, the chances to observe clandestine events improve. Combination of the observations from the four different networks is known as *data fusion*.

2.2 Airborne Radionuclide Monitoring

2.2.1 Particulate Sampling Stations

The 80 particulate sampling stations of the IMS (orange squares in Fig. 2.1) pull air through a filter medium, typically glass or polymer fiber, collecting aerosols that airborne radionuclides are attached to, with the exception of gaseous elements. After 24 hours of sampling, the filter is replaced with a new one and the old one is analyzed with a gamma-ray spectrometer. The short-lived natural radionuclides in the filter are left to decay by starting the radiation measurement of the filter after 4–24 hours. The measurement time is 20–24 hours, typically leading to a 24/24/24 hour cycle denoted as *short-cycle operation*. This operational mode was selected on the basis of the assessment model of Ref. [25] and practical experience.

The most essential requirements on the particulate sampling stations of the IMS are the following: [26]

- air flow $>500\,\mathrm{m}^3$,
- sampling time 24 h,

- decay time $\leq$24 h,
- measurement time $\geq$20 h,
- particulate collection efficiency for filter $\geq$80 % at 0.2 μm diameter,
- particulate collection efficiency for system $\geq$60 % at 10 μm diameter,
- baseline sensitivity $\leq$30 μBq/m^3 for ^{140}Ba,
- data availability $\geq$95 %.

Similar sampling stations are being used by national radiation monitoring networks but usually with less stringent requirements. Typically, the air flow is lower and the work cycle longer (*long-cycle operation*). Thanks partially to the CTBT, more automated sampling stations have become available recently. An example is the Cinderella air sampler of Senya Ltd. [27] that runs air sampling and gamma-ray spectrometry for 15 user-definable working cycles without human intervention.

2.2.2 Analysis of Particulate Measurements

Basic methods of gamma-ray spectrometry are presented in detail in Ch. 4. The aim is to detect *gamma-rays* of various energies from the radioactive decay of radionuclides present in the sample. The individual measured gamma-rays form a spectrum of *gamma-ray peaks* over a wide energy range. These peaks can be analyzed and their properties and associated uncertainties determined using various analysis techniques. From the analysis results it is possible to make deductions about the age and origin of the sample.

The Comprehensive Nuclear Test Ban Treaty requires that the IDC applies "on a routine basis automatic processing methods and interactive human analysis to raw International Monitoring System data" [21, Part I.18], and "the verification activities shall be based on objective information" [21, Article IV.2]. These procedural and quality requirements mean that the analysis results of the verification activities have to be as reliable as possible. In gamma-ray spectrometry, particular emphasis must be placed on radionuclide identification and quantification.

In the CTBT monitoring system, the measured gamma-ray spectra go through an *event screening* process, where they are classified according to the radionuclides present in the sample. Samples containing only natural radionuclides, or well-understood man-made ones, receive low classifications and are automatically excluded from further analysis. Anomalous findings yield high classifications depending on the detected radionuclides. All automatically processed spectra are manually reviewed by human analysts.

Accurate analysis of a typical air filter gamma-ray spectrum, let alone one that contains fresh fission debris, is a demanding task due to the large number of peaks in the spectrum and the uncertainties inherently present in the measurement process. For the results to be reliable, the analysis must be in statistical control throughout. This means, among other things, that all statistically significant gamma-ray peaks are explained. This sets high requirements for the entire radionuclide measurement and analysis system. Within the CTBT network, a high-quality automated gamma-ray spectrum analysis is essential for several reasons:

- Successful event screening requires that fission products are correctly identified. If this cannot be done reliably in the automated routine analysis, the spectrum may not draw the immediate and appropriate attention of the analysts. If the detection of fission products is left for the analysts, there is a higher risk for an evasive sample to pass the review and to be left without a further analysis.
- Event timing is based on ratios of radionuclide concentrations, and thus, it is highly dependent on the accuracy and precision of the spectrum analysis results. An accurate event timing is essential for data fusion with the other monitoring technologies, as well as for the backtracking of plume transport. Large uncertainties in the nuclide concentrations lead to large uncertainties in event timing.
- If a complete interpretation of gamma-ray spectra is the objective, every gamma-ray peak left unidentified in the automatic processing has to be manually analyzed. Thus, reducing the number of false identifications and unexplained gamma-ray peaks speeds up analysis and reduces labor costs.

Analysis of particulate measurements is one of the tasks of the International Data Centre (IDC). Since the IMS stations are operated in a 24 h/24 h/24 h mode, they send a full-time air filter spectrum to the IDC 72 hours after the start of sampling. In addition to the full-time spectra, many stations send preliminary spectra at two-hour intervals. All spectra are processed by the automated analysis pipeline that receives the spectra, stores them in an organized manner, makes a quality control of the spectra, and analyzes them with a tailor-made software that is based on the commercial Genie software package by Canberra [28]. The pipeline has been built on an Oracle database environment.

The preliminary spectra are given closer attention only in an eventual release situation, but under normal circumstances, only the analysis results of the full-time spectra are reviewed by human analysts at the IDC. In addition to the air filter spectra, a 15-minute quality control (QC) spectrum is sent daily from each station to the IDC. The contents of the QC-sources have varied in complexity from ^{60}Co spectra with two major peaks to ^{226}Ra spectra with dozens of peaks. The stations also make periodic background, blank, and calibration measurements that are sent to the IDC and that slightly decrease station availability.

2.2.3 Noble Gas Sampling Stations and Radionuclide Laboratories

Gamma-ray spectrometry of particulate samples is a mature technology with commercial hardware and software packages available. In order to increase the sensitivity of the IMS network, half of the radionuclide stations will be equipped with noble gas sampling. These devices are designed to extract four xenon isotopes from the ambient air, produced in fission processes and not attached to aerosols like all other radionuclides [29].

However, this technology is still under development, illustrated by the fact that there are no certified gas sampling stations in the IMS network yet. Three different radiation measurement methods are applied in the four different prototype gas sampling systems: gamma-ray spectrometry like in particulate sample analysis, beta-gated gamma-ray spectrometry (so-called 2-D coincidence measurements with a single beta-channel and a high-resolution gamma-ray spectrometer), and beta-gamma coincidence measurement (full 3-D coincidence spectrum measurement with scintillation detectors) [29]. None of these systems will be covered in this work.

The IMS radionuclide stations are complemented with 16 radionuclide laboratories (red triangles in Fig. 2.1). One of these laboratories is run by STUK, the Finnish Radiation and Nuclear Safety Authority, which also runs the national monitoring system. Interesting samples can be sent for detailed analysis at the laboratories that can do longer measurements with a gamma-ray spectrometer in laboratory conditions, or sometimes even apply other types of analyses, like mass spectrometry. The methods for gamma-ray spectrometry presented in this work are also applicable in these laboratory analyses.

2.2.4 National Data Centres

According to the CTBT, all State Signatories or national authorities appointed by them have access to the raw measurement data from the IMS network and the analysis results of the IDC. Signatories are entitled to build National Data Centres (NDC) for Treaty verification activities. The NDC's may rely on the analysis results of the IDC or they may run analysis activities on their own. The choice depends on national interests and resources.

For illustration, the Finnish NDC (FiNDC) has been running an automated analysis pipeline for IMS radionuclide data processing since July 1999. The FiNDC pipeline utilizes UniSampo for peak analysis, the newest member of the Sampo family of gamma-ray spectrum analysis codes [16]. It runs under different Unix operating systems, including Linux, Solaris, AIX, and Digital Unix. The core of UniSampo is the same as in previous versions of Sampo, which has been ported to about ten different computer systems in the past [8–15]. This has made it quite portable. The limitations concerning the spectrum size, number of peaks, and fitting intervals have been relaxed as the memory and floating point performance of current Unix workstations are more than sufficient. This has enabled analysis of the most complicated gamma-ray spectra in UniSampo.

The peak analysis results and calibrations from UniSampo are fed to expert system Shaman that is utilized by the FiNDC for nuclide identification [XI, XIV]. The identification results produced by Shaman are used further as input in a spectrum categorizer. The FiNDC pipeline is illustrated in Fig. 2.2. The resources required by this UniSampo–Shaman pipeline are modest: one PC running Linux operating system is sufficient for most users. As a rule of thumb, it takes about 10 seconds to process one spectrum on a modern personal computer, so the theoretical throughput of this kind of a pipeline would be of the order of 8 000 spectra daily. Even with the complete radionuclide monitoring network of the IMS, i.e., 80 stations, the number of full, preliminary, and QC spectra produced daily will be roughly 1 000. Naturally, there is a need for sufficient storage capacity, but with the current hard disk technology this should not be a problem. The data rate from the complete IMS radionuclide network, including spectral data and analysis pipeline results, will be of the order of 500 GB/year with the UniSampo–Shaman system.

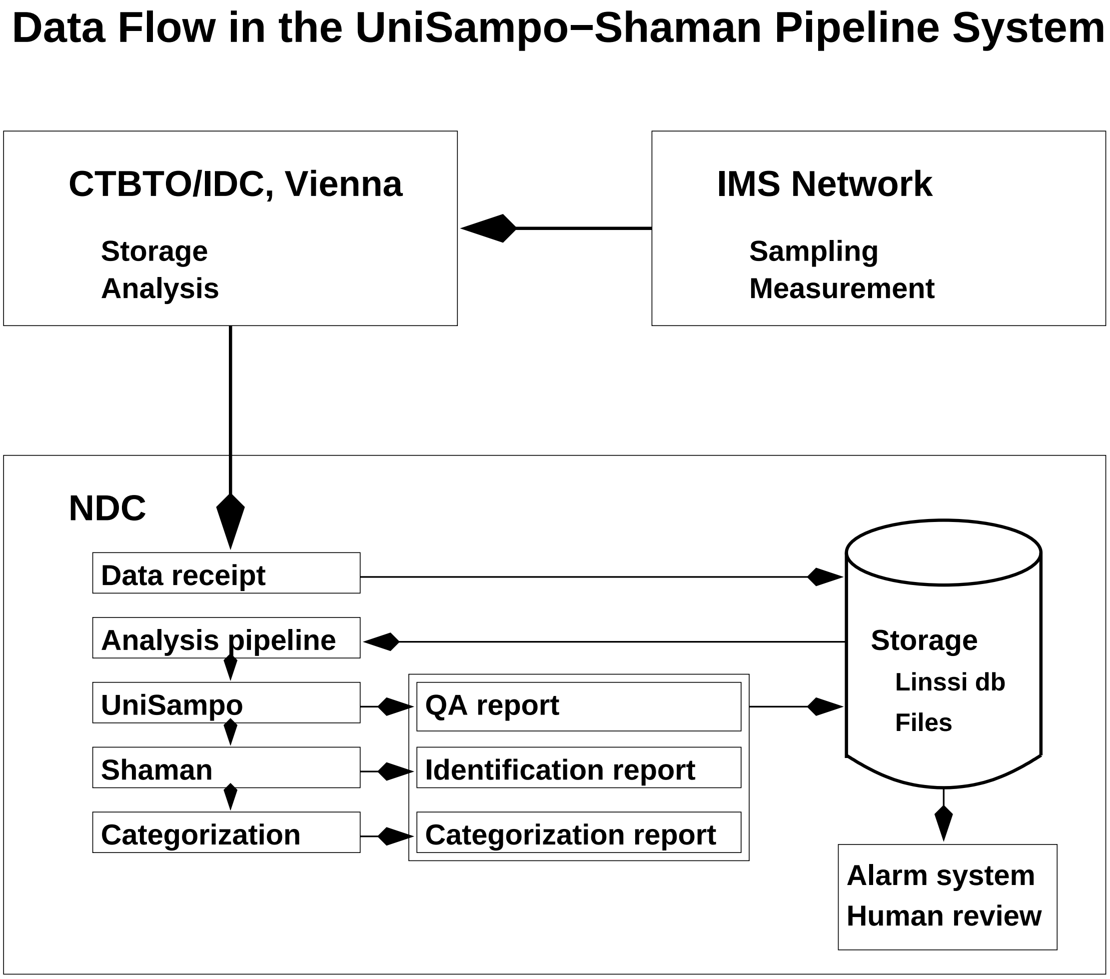

Figure 2.2: Data flow in the FiNDC pipeline.

The computer running the pipeline needs a network connection to the measuring stations. It can be the public internet, a private intranet, or a satellite connection. The basic idea is very simple: spectra are sent by e-mail to a virtual user in the pipeline computer and a script running in the computer polls the mailbox at constant intervals. When a message is received, it is stored in a dedicated directory tree and if the message is a spectrum, it is analyzed with UniSampo and Shaman in batch mode. The analysis results are saved and if they contain anything out of the ordinary, an alarm is sent to an e-mail address. The e-mail alarm can be forwarded to other addresses or sent to cellular phones as a text message if an around-the-clock preparedness is required.

The pipeline operation mode is compatible with an interactive mode where the full-featured graphical user interfaces of UniSampo and Shaman are available to the analyst. By tuning the alert levels, the number of spectra that create an alarm and require an interactive inspection can be adjusted to the level of available analyst resources.

The directory tree of the pipeline works as a primitive database where “queries” can be made with basic Unix commands like `ls` and `grep` and command line wildcards. Our experience is that this kind of a system is sufficient for expert users who know the directory structure well, but a less experienced user would certainly prefer a relational database, especially if dozens of new spectra are processed daily by the pipeline.

Fortunately, open source database systems have recently become available for Linux systems, e.g., `MySQL` [30] and `PostgreSQL` [31]. A collaboration project with TKK, STUK and Health Canada was started in the spring of 2003 to build a relational database for managing the analysis pipeline results. The result of this collaboration is a database schema for gamma-ray spectrometry called *Linssi* [XVIII] that has been made publicly available at `http://linssi.hut.fi`. The current schema bears version number 1.1, but version 2.0 is under development.

Chapter 3

Different Source Terms

A "source term", the quantities of radionuclides released, needs to be defined in order to calculate consequences of a release of any kind. It was first introduced in risk assessment of nuclear power plant accidents in early 1960's [20, p. 250]. When the source term is known by measurement or assumed from theory, transport calculations can be performed for different environmental pathways, accounting for possible mitigation measures. In the first approximation, radiation doses of an exposed population are linearly dependent on the source term if the relative shares of different nuclides remain constant.

The concept of source term can be applied for all radioactive release types. With the methods presented in this work, the aim is to distinguish between different types of anthropogenic releases, always observed in addition to a background of natural radionuclides. The following classification is applied in this work:

1. normal conditions,
2. normal anthropogenic releases, i.e., airborne effluents from operating nuclear power plants (NPP), non-power reactors, spacecraft, accelerators, and hospitals,
3. accident-level releases from these facilities, and
4. releases from nuclear weapon tests.

These different release classes have different relevances for different organizations. However, a common analysis methodology can be applied, at least on the basic level. For all applications, knowledge of the radionuclide contents of the atmosphere in normal conditions, and the daily or seasonal variations of the contents, is a prerequisite to detect the anomalies caused by releases.

3.1 Normal Conditions

In normal conditions there are some 20 radionuclides that are induced from nitrogen, oxygen and argon in the Earth's atmosphere by cosmic rays [20, p. 136]. Of these nuclides, ^{7}Be is produced from nitrogen and oxygen by cosmic ray spallation and it has the highest activity concentration. It is also the only one with a prominent gamma-ray emission, so it is seen in practically all gamma-ray spectra measured from aerosol filters.

The long-lived nuclides ^{232}Th and ^{238}U are abundant in the Earth's crust and soil. A noble gas nuclide belongs to the natural decay chain of each, ^{220}Rn to the thorium chain and ^{222}Rn to the uranium chain. The half-lives of these radon isotopes are 56 seconds and 3.8 days, respectively, and they both diffuse from the ground to the atmosphere. The decay products of these nuclides are not gaseous, so they are attached to aerosols and collected by particulate filters. Decay of ^{220}Rn promptly leads to ^{212}Pb that decays with a 10.6-hour half-life to ^{212}Bi and further to ^{208}Tl. ^{222}Rn on the other hand decays to ^{214}Pb that decays with a 27-minute half-life further to ^{214}Bi and ^{210}Pb. The six radon daughters mentioned here are seen in gamma-ray spectra of aerosol filters.

Even with the best shielding, some background nuclides are always seen in HPGe gamma-ray measurements. One common radionuclide is ^{40}K, a primordial radionuclide with a prominent gamma-ray at 1460.83 keV. This radionuclide that constitutes 0.0117 % of all potassium may also be present in the blank filter, so utilization of both a blank and a background measurement can be recommended. The gamma-ray emitting daughters of ^{232}Th and ^{238}U, as well as the precursor of the ^{235}U chain, are also seen in gamma-ray measurements because of the inevitable thorium and uranium traces in construction materials. The natural chain members most prominently seen in gamma-ray spectra are the following:

- ^{232}Th chain: ^{228}Ac, ^{228}Th, ^{224}Ra, ^{212}Pb, ^{212}Bi, ^{208}Tl,
- ^{235}U chain: ^{235}U only,
- ^{238}U chain: ^{234}Th, ^{234m}Pa, ^{226}Ra, ^{214}Pb, ^{214}Bi, ^{210}Pb.

Since the gamma-emitting nuclides ^{212}Pb and ^{214}Pb and their daughters are present both in the filter and background, the measured activities of these lead isotopes seem to first decay with the exponential decay law but then the decay levels off to the background count rate that is more or less constant. On the basis of their half-lives, ^{214}Pb reaches the background level in 3–5 hours after sampling is ended, whereas it takes 3–5 days for ^{212}Pb to reach its background level.

Cosmic rays also generate peaks that are observable in low-level gamma-ray spectra. Environmental spectra usually belong to this class. The peaks are mainly produced after the fast neutron component of cosmic rays, and the fast neutrons produced by the other cosmic ray components like muons, excite nuclides in the germanium crystal or lead shield and its Cu/Cd/Sn-lining [32–36]. The neutron induced peaks due to inelastic neutron scattering in germanium are very wide and triangular shaped [37], whereas other cosmic ray induced peaks have an ordinary near-Gaussian shape. A list of the most common excitation peaks in air filter spectra with an ordinary shape is given in Table A.1.

In summary, the natural background in gamma-ray spectrometric measurements of airborne radioactivity contains very different components: some of the naturally generated peaks are due to natural radionuclides in the filter, some of them are due to cosmic ray induced excitations in the detector and its neighborhood, some are caused by the filter medium, and some are real background peaks.

3.2 Normal Anthropogenic Releases

On top of the natural background contributions, anthropogenic releases from normal operations can also be classified as one kind of inevitable background. There is a wide variety of man-made releases to the atmosphere, but the most important ones are the following:

NPP and fuel cycle releases: The most abundant radionuclides in gaseous effluents of nuclear power plants are ^{85}Kr and ^{133}Xe [20, p. 242]. There are also other krypton, xenon and iodine isotopes, but they are shorter-lived and decay away in holdup tanks. In the uranium mining process, radon is released to the atmosphere where it blends with the natural concentration. In reprocessing plants, the releases of ^{131}I may be significant if fresh fuel is handled.

Non-power reactor releases: Isotope production reactors produce nuclides for medicine and research purposes. The variety of radionuclides is large and in analogy with power reactors, noble gases and iodine isotopes are the most volatile and thus most prone to escape. Examples of these radionuclides are: ^{125}I, ^{131}I, ^{133}Xe, ^{3}H [38, 39], of which the last one is not a gamma-emitter. The activities are much lower than in power reactors, but may be observable with modern high-volume, high-resolution sampling systems.

Spacecraft releases: Radioisotope thermoelectric generators, commonly known as isotope batteries, are devices that produce electricity from radioactive decay heat. They are used in space applications and remote facilities like lighthouses. The following radionuclides are applied most widely in isotope batteries: ^{90}Sr, ^{210}Po, ^{238}Pu [20]. Especially ^{238}Pu has been used in spacecraft and it has spread in the atmosphere when satellites have re-entered the atmosphere. All radionuclides utilized in these devices are weak gamma-emitters to minimize the need for shielding. This is in contrast to a Soviet satellite powered by a nuclear reactor that crashed in Canada in 1978. It spread a variety of fission and activation products, in addition to its uranium fuel [20, p. 401].

Accelerator releases: Cyclotrons and other particle accelerators are also used for radionuclide production purposes. Again the gaseous and iodine isotopes are most prone to escape, e.g., ^{13}N, ^{15}O, ^{123}I [38, 39]. However, the production processes vary and some radionuclides are produced as decay products of gaseous radionuclides and may thus escape in an intermediate phase. In high-energy proton-proton accelerators, virtually any nuclides are generated but in small quantities [XV]. In analogy with non-power reactor releases, the releases from accelerators are low but observable with modern sampling systems.

Hospital releases: Radionuclides are used either for diagnostic or therapeutic reasons at hospitals. These radionuclides are usually short-lived, so the wastes containing them are diluted and released in normal sewage or incinerated and thus released to the atmosphere [20]. Air sampling systems readily detect the radionuclides that are utilized most widely, e.g., ^{99}Mo/^{99m}Tc, ^{123}I, ^{125}I, ^{131}I. The volatile fission products among these nuclides like ^{131}I are indistinguishable from NPP releases.

It can be seen that the variety of radionuclides observed in a normal situation can be quite large. When introducing a radionuclide monitoring system, either for a single facility, for a territory, or for global surveillance like the International Monitoring System for the CTBT, this has to be taken into account. Baseline data needs to be obtained for a sufficiently long period of time, typically 1 year or longer [20], in order to distinguish the anomalous observations among the normal ones.

3.3 Accident-Level Releases

Composition of the core of a nuclear reactor changes over time. Concentrations of radionuclides can be estimated with isotope generation and depletion codes like, e.g., ORIGEN-2.2 [40, 41]. However, the step from the core composition to the actual source term in a severe accident situation is not a trivial one. The regulations in the United States state that the accident release to the reactor building should be calculated as 100 % of the noble gases, 50 % of the iodines, and 1 % of the remaining fission products, and one percent per day of the reactor building contents would leak to the outside atmosphere [20, p. 250]. These fractions have been shown to overestimate the releases, but they are still the basis for the source term in NPP risk analyses.

The vast majority of the radionuclide inventory are fission products. In the comprehensive fission yield compilation of Ref. [42] there are over 800 radionuclides with a non-zero cumulative fission yield. The smallest yields are of the order of 10^{-14} and the largest ones almost 10 %. If the stable nuclides and short-lived ones ($T_{1/2} < 26\,\mathrm{min}$) are omitted, the list reduces to 44 fission products according to the U.S. Nuclear Regulatory Commission as cited in [20, p. 248]. Of these nuclides, the I, Te, and Cs isotopes are most relevant for dose estimates and they are also most prominent in measured gamma-ray spectra of air filter samples. In contrast, the pure beta-emitters on the list like ^{89}Sr and ^{90}Sr are not seen in gamma-ray spectra at all, and the Kr and Xe isotopes are seen only if they are trapped in a fuel matrix particle attached to aerosols.

In addition to fission products, nuclear reactors produce activation products from the construction materials, mainly steel, and from the fuel through parasitic neutron captures. The most significant activation products, according to the U.S. NRC list cited above, are the cobalt isotopes ^{58}Co and ^{60}Co, both strong gamma-emitters, and the following transuranic nuclides: ^{239}Np, ^{238}Pu, ^{239}Pu, ^{240}Pu, ^{241}Pu, ^{241}Am, ^{242}Cm, and ^{244}Cm. Only ^{239}Np and ^{241}Am have sufficiently strong gamma-ray emissions to be observed in aerosol samples.

In this work, gamma-ray spectra of air filters sampled from fresh reactor releases have been analyzed. The findings are consistent with the general principles above, but there are some additional radionuclides observed that are not listed here. Sec. 6.3 presents the results and discussions.

3.4 Releases from Nuclear Weapon Tests

Nuclear weapons are devices that release vast amounts of nuclear energy in a short period of time. These devices utilize the fission and fusion processes in different combinations: the first nuclear weapons utilized only fission, the second generation were fission devices boosted with a fusion core, the third generation were thermonuclear weapons with a fission trigger and a fusion device, and the fourth generation were fission-fusion-fission weapons [43, 44]. The maximum yield of the weapons has increased from one generation to the next one.

The only radioactivity released from a fusion process would be ^{3}H, but the fusion neutrons generate a variety of activation products near the site of detonation. However, all generations of nuclear weapons utilize fission devices that release fission products. They are most readily observed with gamma-ray spectrometers. Fission neutrons also generate activation products from the construction materials and the surrounding bulk material, be it air, seawater, soil, or rock.

In the first approximation, the source term from nuclear test releases is very similar to that of NPP releases. There are differences like the harder neutron spectrum (comparing to thermal reactors that are the majority) and lack of release barriers (fuel matrix, cladding, pressure vessel, containment) in a test release, but their effect on the observations is difficult to quantify. There are, however, some inherent differences between the two that can be applied to infer the release type:

1. In a weapon all fission products are generated at the same time that is typically within a few days of the observation, whereas fission products accumulate in a reactor over a period of 3–4 years. Therefore, the decay chains in a weapon release may not have reached the equilibrium. Useful decay chains for these studies are the following: ^{99}Mo–^{99m}Tc that reaches 90 % of the equilibrium in 1 day, ^{140}Ba–^{140}La that reaches 90 % of the equilibrium in 6 days, and ^{95}Zr–^{95}Nb that reaches 90 % of the equilibrium 250 days. (These figures are based on simple decay chain calculations and thermal ^{235}U fission yield data of Ref. [42].) A similar reasoning can be used in noble gas sampling systems for the nuclide pair ^{133m}Xe–^{133}Xe [45].

2. Since the neutron chain reaction in a weapon is much shorter (of the order of 1 μs [44]) than in a reactor, a weapon produces less activation products in proportion to fission products than a reactor. The most common example is ^{134}Cs whose fission yield is practically zero. It is produced by neutron activation from ^{133}Cs, the stable nuclide ending the decay chain of fission products with mass number 133 [20]. If the concentration of ^{134}Cs is comparable to that of ^{137}Cs that is an abundant fission product, the release is likely to come from a NPP, otherwise from a weapon test.

   By inspection of the fission product chains in Ref. [46], ten nuclides analogous to ^{134}Cs and with a reasonable half-life for detection can be found: ^{82}Br, ^{86}Rb, ^{94}Nb, ^{110m}Ag, ^{122}Sb, ^{124}Sb, ^{130}I, ^{142}Pr, ^{152}Eu, ^{154}Eu. These nuclides are more typical in a reactor release than in a weapon release, but of course, their concentrations in a weapon test release depend on the element composition of the test site.

3. The activity ratio of the abundant and long-lived fission products ^{137}Cs and ^{90}Sr can be used to infer the nuclide that has fissioned. This ratio is inherently different for ^{235}U, ^{238}U and ^{239}Pu: 1.09, 1.78 and 3.07, respectively, assuming a fission neutron spectrum [44, p. 91]. Current NPP's utilize uranium fuel or in some countries so-called mixed-oxide (MOX) fuel with part of the ^{235}U replaced with ^{239}Pu, so the ^{137}Cs/^{90}Sr ratio should be somewhere between 1 and 2 in reactor releases. (Unfortunately, ^{90}Sr

and its shorter-lived daughter ^{90}Y are pure beta-emitters, so they cannot be measured by gamma-ray spectrometry.)

4. The fissile uranium isotope ^{233}U that can be made of ^{232}Th in a breeder reactor is not currently utilized in NPP's. If a bomb were made of ^{233}U, it would inevitably contain ^{232}U that is also produced from ^{232}Th. This uranium isotope would feed the natural ^{232}Th decay chain bypassing ^{228}Ac, thus generating a situation where ^{224}Ra would be in excess relative to ^{228}Ac [47]. If this kind of a nonequilibrium is observed, it hints to a ^{233}U weapon test.

5. The ratio of the fuel activation products ^{239}Np and ^{237}U is a measure of the thermonuclear character of a weapon test [47]. According to past experience, pure fission bombs have ^{239}Np/^{237}U ratios above 30, whereas the ratio is around unity for thermonuclear weapons.

6. The isotopic composition of plutonium produced by nuclear reactors is different from that in nuclear weapons, so the isotopic ratio ^{240}Pu/^{239}Pu can be used to distinguish the source of plutonium [20, 48]. (Unfortunately, these Pu isotopes are weak gamma-emitters and their alpha spectra are very similar, so they can be distinguished only with mass spectrometry.)

These and other similar criteria can be applied to identify the source of an anomalous release of fission products. The more criteria are applicable, the more reliable is the conclusion. Obviously, a prerequisite is that the radionuclides are identified correctly and their activities are determined accurately on the basis of the measured gamma-ray spectrum.

Chapter 4

Gamma-Ray Spectrum Analysis Methods

In this Chapter, methods applicable to gamma-ray spectrometry are presented. The emphasis is on some novel concepts that have been found to work reasonably well and have been implemented in expert system SHAMAN (Ch. 5). We start with the fundamentals of gamma-ray spectrum analysis that are necessary for the remainder of this work, and then proceed to calibration considerations, limits of detection, quantitative prediction of special peak areas, and decay corrections.

Our goal is not to reproduce the contents of excellent text books on the subject, including Refs. [22], [49], and [50]. In general, we will not review all existing analysis methods nor present a critical evaluation of the methods.

4.1 Characteristics of Gamma-Ray Spectra

Gamma-rays are quanta of electromagnetic radiation, i.e., photons, produced when an atomic nucleus changes its energy state from a higher state to a lower one. This is in contrast to X-rays that are produced by similar de-excitations in the electron energy states. The energies of gamma-rays span to higher energies than those of X-rays, but the energy regions overlap in the lower end, below 150 keV. Even if the difference in the genesis of gamma-ray and X-ray photons is crucial, they cannot be distinguished from each other upon detection, at least generally, in a trivial manner. X-ray peaks are an integral part of a measured gamma-ray spectrum and they need to be accounted for when analyzing the spectrum. In many instances below, the term *gamma-ray* can be implicitly replaced with *gamma-ray or X-ray*.

The term *nuclide* is used for a category of atomic nuclei, namely those having a certain number of protons and neutrons. Some nuclides are stable like $^{59}_{27}\mathrm{Co}$, but a great majority of known nuclides are radioactive like $^{60}_{27}\mathrm{Co}$, hence called *radionuclides*. Some nuclides have *isomeric states*, which are excited states with a significant half-life. For example, $^{60}_{27}\mathrm{Co}$ and $^{60m}_{27}\mathrm{Co}$ represent the ground and isomeric states with half-lives 5.27 a and 10.5 min, respectively, the m-state lying 58.6 keV above the ground level. The distinction between a

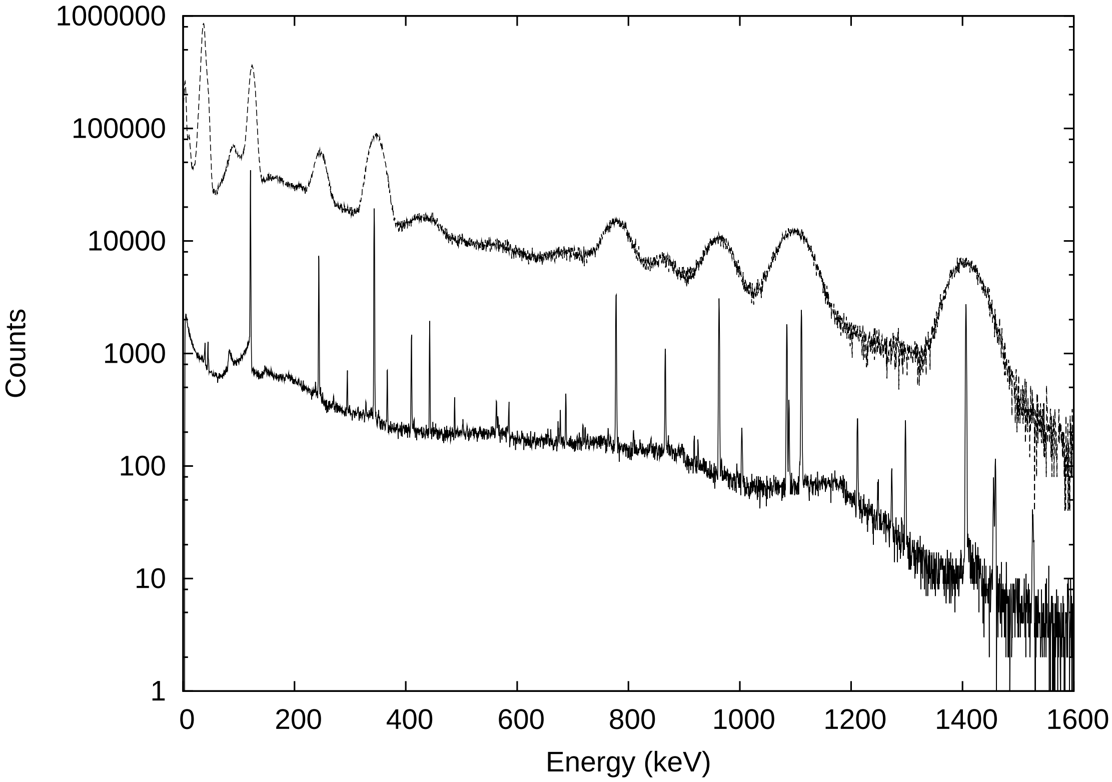

Figure 4.1: A comparison of two ^{152}Eu spectra: one measured with a Ge-detector and the other synthesized imitating a NaI(Tl)-detector. The difference in energy resolution is substantial: some 12 peaks can be distinguished from the NaI(Tl)-spectrum, 40 peaks from the Ge-spectrum. (The NaI(Tl)-spectrum has been shifted upwards to avoid overlap.)

nuclear state that is considered isomeric and one that is not is a matter of definition, if the state only de-excites to the lower-lying states, but if an excited state can decay by α, β or other decay mode, then it is an isomeric state regardless its half-life. The shortest-lived nuclear states considered as isomers are typically in the nanosecond range. Throughout this work, no distinction is made between ground and isomeric decaying states.

An excited nuclide, generated as a result of radioactive decay or inelastic scattering, de-excites by two competing processes: gamma-ray or internal conversion electron emission. This typically occurs in a matter of picoseconds and leads to another state that may be stable or that may decay further. Since the energy states of a nuclide are discrete, the emitted gamma-rays and electrons belong to one or several monoenergetic groups. The energy states and thus, gamma-ray and electron energies are nuclide-specific, enabling different radionuclides to be distinguished from each other. The conversion electron spectrum can be measured for this purpose, but gamma-rays are always less disturbed by eventual absorbers and usually more abundant, so they are typically chosen for energy spectrum measurements.

The detector types applicable to gamma-ray spectrum measurements are different scintillation detectors, e.g., NaI(Tl), CsI(Tl) or BGO, and semiconductor detectors, e.g., Si, CdZnTe or Ge. The measuring efficiency of gas-filled detectors, on the other hand, is usually insufficient for gamma-ray spectrometry purposes. An advantage of scintillation detectors is that they can be operated at ambient temperature, whereas a Ge-detector must be cooled down to suppress electronic noise. However, scintillation detectors have a relatively poor energy resolution, typically 40 keV at 662 keV gamma-ray energy [22], which limits their application to spectra with only a few gamma-ray peaks. A comparison of a ^{152}Eu-spectrum measured on a Ge-detector with 16.5 % relative efficiency and a synthesized NaI(Tl)-spectrum in Fig. 4.1 reveals a substantial difference in energy resolution. This is why Ge-detectors are preferred over scintillation detectors in all but the simplest applications.

Silicon semiconductor detectors of sufficient size for detecting gamma-rays can be manufactured only with the lithium-drifting technology, previously used also for Ge-detectors, because the purity available for silicon is lower than for germanium [22]. Still, the lower atomic number of silicon makes it less efficient for gamma-rays, so silicon detectors are mainly applied for charged particles and X-ray spectrometry.

Several other semiconductor materials besides silicon and germanium have been examined over the years, e.g., CdTe, GaAs and HgI_2. Currently, cadmium zinc telluride (CdZnTe or CZT) is the third most common choice in photon spectrometry. The advantage of CdZnTe detectors is that they can be operated at room temperature, but the disadvantage is a worse energy resolution: the best CdZnTe detectors achieve a resolution of 10 keV at the 662 keV gamma-ray energy of ^{137}Cs [51]. This is much better than scintillation detectors, but still much worse than current Ge-detectors that routinely have a resolution of 1.5 keV at this energy. Additionally, the CdZnTe detectors are available in smaller sizes than Ge-detectors, typically $10 \times 10 \times 10\,\mathrm{mm}^3$, resulting in smaller measuring efficiencies. In some applications, pros of CdZnTe may outweigh cons, but they cannot replace Ge-detectors when complex gamma-ray spectra are measured. *From this point on, this work will concentrate solely on gamma-ray spectrometry with Ge-detectors.*

A Ge-detector converts the line spectrum emitted by a radionuclide into a measured gamma-ray spectrum where distinctive peaks are seen on a relatively smooth baseline. All essential information on the source is carried by the peaks: the positions are proportional to original gamma-ray energies and areas to nuclide activities. Low-resolution gamma-ray spectra are typically analyzed using spectrum deconvolution or unfolding where the response function of the measurement setup is assumed to be known [22, p.672]. In high-resolution gamma-ray spectrometry, on the other hand, only the peak shape function needs to be modeled. It is much less dependent on the measurement setup than the complete response function.

The structure of a spectrum reflects the interactions of gamma-rays in the detector material leading to the detection of a gamma-ray photon. The three principal interaction mechanisms, photoelectric absorption, Compton effect and pair production, produce different features in the spectrum, and they will be explained briefly in the following. The synthesized spectrum of Fig. 4.2 illustrates these features [IV].

Photoelectric absorption occurs when a gamma-ray interacts with a bound electron, transferring all its energy to the electron and disappearing itself. It leads to a full absorption of the gamma-ray energy to the detector. When a large number of monoenergetic gamma-rays experience photoelectric absorption, a peak is created in the pulse height spectrum. This peak is called a *full-energy peak* or a *photopeak*, and its shape is near Gaussian in a well-behaved spectrometer. The finite width of the photopeak is in practice completely due to the finite energy resolution of the Ge-detector and signal-processing electronics, since the

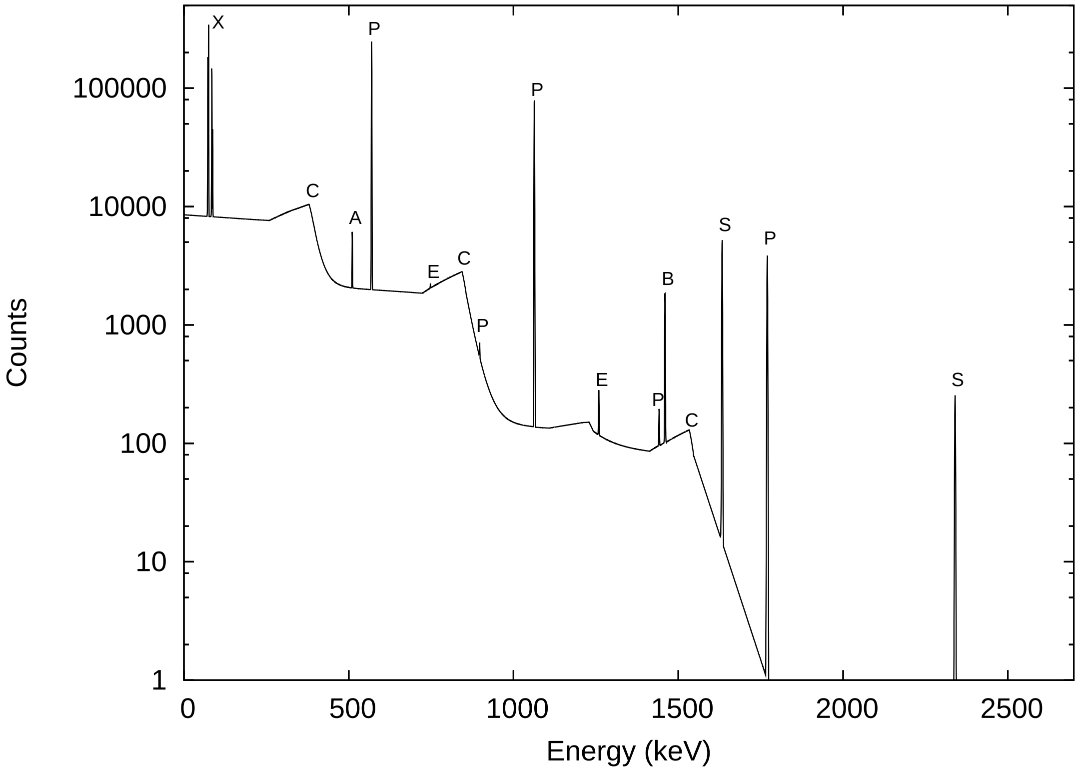

Figure 4.2: A synthesized spectrum of ^{207}Bi [IV]. The statistical fluctuations have been omitted from the figure on purpose, in order to show the different spectrum features. This is why the spectrum looks unnatural. Peaks at the following energies (in keV) are seen in the spectrum: 72.8, 75.0, 84.8 and 87.3 (X-ray multiplet = X), 511.0 (annihilation peak = A), 569.7 (ordinary photopeak = P), 748.2 (double escape peak = E), 897.8 (P), 1063.7 (P), 1259.2 (single escape peak = E), 1442.2 (P), 1460.8 (background peak = B), 1633.4 (sum peak = S), 1770.2 (P) and 2339.9 (S). Clear Compton edges (C) are seen at the energies 393, 858 and 1547 keV.[a]

natural line width of gamma-rays is far smaller than it. In Fig. 4.2, the most prominent photopeaks are seen at the major gamma-ray energies of ^{207}Bi: 570, 1064 and 1770 keV.

Compton effect means elastic scattering of a gamma-ray from an electron. Any fraction of the gamma-ray energy can be transferred to the electron in the interaction, because all scattering angles are allowed. The maximum energy of a Compton electron corresponds to a scattering angle of 180°. The gamma-rays that interact through Compton scattering produce a relatively smooth *Compton continuum* in the spectrum, spanning from zero energy up to a *Compton edge* at the maximum energy of the Compton electron. In Fig. 4.2, the clearest Compton edges are seen at the energies 393, 858 and 1547 keV, corresponding to the most prominent gamma-ray energies mentioned above.

---

[a] The spectrum has been synthesized with SYNTE, a non-commercial extension of SAMPO [17–19]. The calibrations and time parameters have been taken from the real spectrum in Fig. 4.3. The effect of coincidence summing (Sec. 4.7) has been taken into account in the photopeak areas and the sum peak areas have also been obtained from this calculation. The escape peak areas have been determined with the help of escape peaks in the real spectrum. The areas of the annihilation and background peaks have been chosen arbitrarily.

A gamma-ray may experience multiple interactions in the detector, and since their time interval is of the order of 10^{-10} s, far below the resolving time of a typical Ge-spectrometer, energy absorptions in subsequent interactions sum up to a single pulse. Thus, if a Compton scattered gamma-ray experiences photoelectric absorption within the detector, a pulse is not added to the Compton continuum but to the full-energy peak. Similarly, two or several subsequent Compton scatterings may lead to a pulse in the *multiple Compton effect region* between a Compton edge and the corresponding photopeak. This is why the Compton edges are not completely vertical as seen in Fig. 4.2.

If the energy of a gamma-ray exceeds the rest mass of two electrons or 1022 keV, it may interact through pair production in the detector. In pair production, possible only in the vicinity of a nucleus, the gamma-ray disappears and an electron-positron pair is created. The gamma-ray energy in excess of 1022 keV appears as kinetic energy of the pair. This energy is usually absorbed in the detector as the range of the particles is less than 1 mm at typical energies. Usually the positron reaches a thermal energy before it annihilates with an electron, thus creating two 511 keV gamma-rays known as annihilation quanta.

The annihilation quanta, like gamma-rays originating from outside of the detector, can interact in the detector through photoelectric absorption or Compton scattering, or they can escape from the active volume of the detector. If both annihilation gamma-rays escape without further interactions, the original gamma-ray conveys an energy of $E_\gamma - 1022$ keV to the detector. From a large number of such events, a *double escape peak* may be produced in the spectrum 1022 keV below the corresponding photopeak. Similarly, if one annihilation gamma-ray escapes and the other is totally absorbed in the detector, an energy of $E_\gamma - 511$ keV is conveyed to the detector and these events may produce a *single escape peak* in the spectrum. In the case of ^{207}Bi, the only gamma-ray with a significant pair production probability is the one at energy 1770 keV. The corresponding escape peaks are observable at energies 748 and 1259 keV in Fig. 4.2.

Annihilation gamma-rays are produced nearby the source, if it contains β^+-decaying nuclides, and in the structures around the spectrometer if gamma-rays above 1022 keV are abundant, thus experiencing pair production and subsequently producing annihilation gamma-rays. These gamma-rays often produce an *annihilation peak* at the energy 511 keV. The annihilation peak cannot be used for nuclide identification purposes, because there are several hundred β^+-decaying nuclides that have the 511 keV line on their list of gamma-ray lines. Therefore, if a β^+-decaying nuclide like ^{36}Cl and ^{59}Ni has no other gamma-ray energies, it cannot be reliably identified by gamma-ray spectrometry, but these nuclides are rare.

Sum pulses are produced when two or more gamma-rays deposit all or part of their energy to the detector within the resolving time of the spectrometer. If simultaneous full energy absorptions occur frequently, *sum peaks* are produced in the spectrum. Summing happens under two distinct circumstances: if the count rate in the detector is high (random summing) or if nuclides with complicated decay schemes are measured in a close geometry (true coincidence summing).

Random sum peaks are created at the sum energies of the most abundant gamma-rays emitted by the source. The counting rate of random sum pulses is proportional to the product of the counting rates of the summing gamma-rays. Random sum peaks are often wider than ordinary photopeaks and the lower side tail tends to be stronger due to partially summed pulses. Furthermore, partially summed pulses contribute to the continuum at the low energy side of the sum peak, thus producing a significant step under the peak.

True coincidence sum peaks result from a detection of two gamma-rays that have been emitted in a cascade from the same decaying nuclide, with a time interval shorter than the resolving time. The existence of coincidence sum peaks indicates a high measuring efficiency, because the probability of coincidence summing is proportional to the solid angle subtended by the detector, i.e., coincidence summing is significant at close measuring geometries regardless of the counting rate. In contrast to random sum peaks, coincidence sum peak areas are linearly dependent on the source activity. The shape of coincidence sum peaks is usually close to that of ordinary photopeaks, but if both random and coincidence summing occurs, sum peaks may be wider than usual. In Fig. 4.2, there are coincidence sum peaks at the energies 1633 and 2340 keV that result from cascading gamma-rays $1063 + 570$ keV and $1770 + 570$ keV, respectively.

As stated above, X-rays are produced in the transitions between electron states of an atom. In nuclear decay, excited electron states are created in internal conversion and electron capture processes. If these processes are abundant in the source, the full-energy peaks of X-rays or *X-ray peaks* are seen in gamma-ray spectra, especially if the source contains heavier elements since probability of internal conversion increases with increasing element number Z. In principle, X-ray peaks have a different shape than gamma-ray peaks at the same energy due to their larger natural line width, but the difference is not always detectable in measured spectra. The observed peak width is a convolution of the line width and the finite energy resolution of the spectrometer. X-ray peaks are only element-specific and cannot generally be used alone in nuclide identification, but their sizes relative to the gamma-rays are tabulated sufficiently accurately to be useful in confirming a nuclide identification. If the set of candidate nuclides can be limited by some means (like in noble gas sampling), then X-rays may be sufficient for identifying nuclides by themselves. In Fig 4.2, the KX-ray multiplet of lead between 70 and 90 keV is observed, because ^{207}Bi decays almost solely via electron capture to ^{207}Pb, accompanied with X-ray emissions.

Construction materials of the spectrometer and the surrounding structures contain radioactive nuclides. Typical observations in gamma-ray spectra are ^{40}K, having a share of 0.0117 % all potassium, and the members of the natural decay chains originating from ^{232}Th, ^{235}U and ^{238}U that create a number of *background peaks* in measured spectra. In low-level measurements, these peaks are minimized by efficiently shielding the spectrometer with suitable materials. Another important source of background peaks is cosmic radiation: especially its muon and fast neutron components produce excitations in the Ge-detector and its shielding that de-excite by gamma-ray emission. This background component is extremely difficult to avoid in low-level measurements. In Fig 4.2, there is a background peak at 1461 keV, the characteristic ^{40}K energy.

A feature of a measured gamma-ray spectrum not illustrated in Fig. 4.2 are the *X-ray escape peaks*. They are a result of germanium X-rays escaping from the active volume of the detector, leading to an energy transfer deficit of 10 keV. A large number of such escapes lead to X-ray escape peaks 10 keV below each large peak. This phenomenon is strongest at energies below 100 keV, but some modern detectors generate X-ray escape peaks even for peaks above 1000 keV. However, X-ray escape peaks at high energies are 3–4 orders of

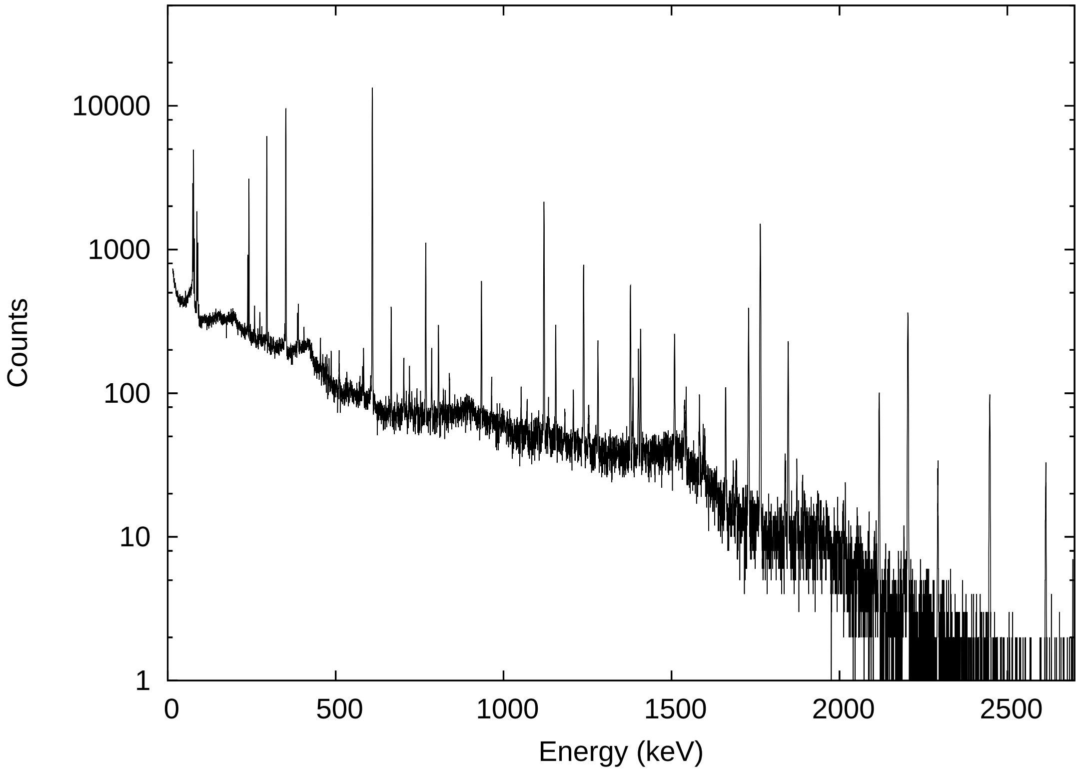

Figure 4.3: A measured air filter spectrum. Approximately 75 peaks can be found from this spectrum, including photopeaks, escape peaks, an annihilation peak, sum peaks, and X-ray peaks.

magnitude smaller than the corresponding photopeak, making their detection difficult.

Another feature of a measured gamma-ray spectrum omitted from Fig. 4.2 is the *backscatter peak*. It is produced by the gamma-rays which have Compton scattered to the detector from the materials behind it, thus experiencing a scattering of approximately 180°. The backscatter peak is asymmetrical and relatively wide, ranging from the energy corresponding to a scattering of 180° upwards. It can be distinguished from the other peaks, usually automatically, in the same way as the Compton edges as it resembles an inverse Compton edge.

In real measured spectra, there usually exist dozens of peaks and the recognition of the different spectrum components becomes difficult. Additionally, statistical channel-to-channel fluctuations contribute to the difficulties. A typical air filter spectrum in a release situation is shown in Fig. 4.3.

4.2 Analysis of Gamma-Ray Spectra

In general, the aim of gamma-ray spectrometry is to identify the nuclide contents of the source and to determine nuclide activities on the basis of the peaks found in the spectrum. The sophistication of the task varies substantially from one application to another. In simplest cases, the present nuclides are known in advance and their activities are determined by comparison to suitable standard spectra, whereas the most complex gamma-ray spectra may include hundreds of peaks originating from totally unknown nuclide contents. A complete analysis in the complex cases requires a comprehensive reference library and several hours, even days, of expert labor.

The main phases during gamma-ray spectrum analysis are:

1. peak search,
2. peak area determination,
3. nuclide identification, and
4. activity determination.

It is relatively straightforward to automate the analysis phases 1 and 2 as can already be seen from the number of published gamma-ray spectrum analysis programs, some of which are presented in Refs. [8] – [16] and [52] – [84]. Many of these programs also implement the analysis phases 3 and 4, but mostly in a limited manner. This is because the nuclide identification phase combines mathematical calculations with numerous empirical conclusions that are difficult to implement in a computer program in a general manner.

All essential information in a high-resolution gamma-ray spectrum is included in peaks that stand on a more or less smooth baseline. However, finding peaks is not a trivial task in a multichannel spectrum where the pulse counts fluctuate from one channel to another according to Poisson statistics. Several peak search algorithms have been implemented, the main principles being smoothing, differentiating or direct filtering of the spectrum. In the Sampo program, for example, peak search is based on a method of generalized second differences presented by Mariscotti [85]. Peak search is usually done on the raw measured spectrum, but it is also possible to subtract the continuum before peak search [86, 87]. The result of the peak search phase is a list of peak centroids that can be converted into energy units with the help of an *energy calibration*.

A basic feature of any peak search method is an adjustable sensitivity level: the lower the level, the more peaks are found, both real and spurious ones. Adjustment of the sensitivity level is a matter of trade-off between the number of spurious peaks, also known as Type I errors of peak search, and the number of unfound real peaks, also known as Type II errors. The optimum is application-specific.

If the energy calibration is sufficiently reliable or if it can be inferred from the measured spectrum, a library-oriented peak search can be applied [87]. This method is actually not a peak search, since peaks are set at listed, pre-selected energies and are tested against a critical limit. However, the result is a list of peak energies like from peak search.

There are two main methods for peak area determination: summing of counts above the baseline in the vicinity of a peak or fitting analytical peak and baseline functions to the spectrum and determination of the peak area by integration. The latter method is more reliable, especially when determining areas of multiplet components or small peaks. Several different peak shape functions have been implemented as listed in Ref. [49, p. 164–165]. They are typically composed of a Gaussian main component, some tailing terms and a step function underneath the Gaussian peak. Especially some multiplet components may be impossible to resolve in the peak search phase, but they can sometimes be recognized in the peak area determination phase. Some spurious peaks can also be discarded in this phase, especially if peak areas are determined in an interactive session. The determined peak areas can be converted into gamma-ray emission rates with the help of a *peak efficiency calibration*.

The analysis phases described above produce a list of gamma-ray energies and emission rates. The aim in the nuclide identification phase is to explain the observed peaks with the gamma-ray lines of a reference library, taking into account statistical uncertainties in peak parameters and reference data. A strict algorithm cannot be presented for the nuclide identification phase, but it proceeds roughly in the following manner (slightly modified from Ref. [88]):

1. Take the unexplained peak with the highest energy, because there are fewer candidate nuclides for the higher energy gamma-rays in the reference library. Take the nuclides with gamma-ray energies within a chosen tolerance of the peak energy as possible candidates.

2. Check if the other gamma-ray lines, especially the more prominent ones, of each candidate nuclide can be associated with the spectrum peaks. If not, discard the candidate.

3. Check that the half-life of each candidate is not too short compared to the waiting time between receiving and measuring the sample. If a nuclide has a longer-lived parent nuclide, use its half-life as the effective half-life, since the daughter may have been produced in the decays of the parent during the sampling/irradiation and waiting time. Discard candidates with too short an effective half-life.

4. If the origin of the sample is known and a candidate or any of its parents cannot be produced with a conformable reaction, discard the candidate.

5. Check the correspondence of peak and line proportions of each candidate. If the proportions match within a tolerance, consider the peaks as explained by this nuclide alone. Other candidates must be found for underexplained peaks.

6. If there are unexplained peaks, which cannot be explained as X-ray, sum or escape peaks of the accepted nuclides, return to phase 1.

When a set of suitable candidate nuclides has been selected by applying the strategy above, an activity calculation can be tried. There are two different activity estimates: a *primary line activity* is computed on the basis of the most prominent gamma-ray peak of the nuclide and a *least-squares activity* is calculated on the basis of all gamma-ray peaks of the nuclide. If uncertainty estimates, most importantly those of peak areas and efficiencies, are under statistical control and if the set of candidate nuclides is complete, the latter method should yield a more reliable estimate. Generally, the first set of nuclides is not the correct one, which may be seen from the resulting activities, and one must discard some of the candidates and possibly reconsider some of the previously discarded ones. Several iterations are usually required, because it is difficult to select the most unfitting candidates. It requires some expertise to decide when the peaks have been explained sufficiently well to cease the iterations.

The result of gamma-ray spectrum analysis should be the most probable nuclide composition of the sample and best estimates for their activity values. The uncertainties of the calculated activities should also be estimated, taking into account all uncertainty components, the peak area and efficiency uncertainties being the most prominent ones. If the nuclide identification is reliable, it is possible to draw further conclusions concerning the sample. In activation analysis the elemental composition of the sample is of interest, and in environmental monitoring anthropogenic nuclides are searched for, because they can indicate a radioactive release. This kind of conclusions are a task for an experienced analyst or a sophisticated expert system.

4.3 Calibration Considerations

In the discussion above, it was assumed that the spectrometer had been properly calibrated prior to starting analysis of a sample of interest. Calibrations would deserve a thorough discussion, since the reliability of spectrum analysis and nuclide identification results is heavily dependent on the quality of calibrations. For example, an erroneous shape calibration may lead to unfound peaks in the peak search phase and an erroneous efficiency calibration to erroneous nuclide activities. However, only some reflections on essential calibration issues are presented in this work.

The *energy calibration* and *peak efficiency calibration* are necessary for converting spectrum channels to gamma-ray energies and peak areas to gamma-ray emission rates and further to activities, respectively. Additionally, a *peak shape calibration* and *total efficiency calibration* are used by some analysis methods. The former may be utilized for defining a suitable width parameter for peak search and for peak area determination with pre-defined peak shapes like in the SAMPO program. The latter is required when calculating coincidence summing corrections (Sec. 4.7).

4.3.1 Energy Calibration

The channel-energy relation is very linear with modern Ge-spectrometers, especially when using a digital multichannel analyzer (MCA). Therefore, either linear interpolation of calibration data pairs or a straight line fitted to them usually works well and these functions are robust enough for automated pipeline processing. Deviations below 0.5 keV are typically achieved between data points and the fitted line. In interactive analysis, a third-degree polynomial may be the optimum selection, but the increase in accuracy is counterbalanced by a decrease in robustness, so this choice should be used with caution in pipeline processing.

Even if the energy calibration remains almost linear, the channel-energy response of a spectrometer will inevitably drift over time. If frequent recalibration is not practical, the drift can be compensated with hardware, i.e., a spectrum stabilizer, or with software, i.e., an energy calibration update also known as an internal energy calibration. The software solution can be automated, if each measured spectrum contains some characteristic peaks like air filter spectra do, but in any case it can be done in interactive analysis. By matching the spectrum peaks with known gamma-ray energies, an energy calibration sufficiently accurate for nuclide identification is achieved, i.e., peak energies remain within 0.5 keV of the correct values.

This kind of an internal energy calibration has been implemented both in the pipeline analysis software of the International Data Centre (Sec. 2.2.2) and that of the Finnish National Data Centre (Sec. 2.2.4). For illustration, the latter applies a simple two-pass method:

1. Search peaks with search threshold 7.0 (relatively high).
2. Match the found peaks with those in a characteristic peak list using a tolerance of 4.0 keV and the nominal energy calibration.
3. Calculate a preliminary energy calibration by fitting a first or third degree polynomial to the channel-energy pairs.
4. Search peaks with search threshold 4.0 (medium).
5. Match the found peaks with those in a characteristic peak list using a tolerance of 1.0 keV and the preliminary internal energy calibration.
6. Calculate the final energy calibration by fitting a 1st or 3rd degree polynomial to the channel-energy pairs.

According to our tests with a large set of air filter spectra, this two-pass method works reliably. It typically brings peak energies to within 0.1 keV of the library values. The method was also tested with two sets of air filter spectra measured in release situations and it succeeded in all of these cases having a larger number of peaks than usually. The method seems to fail only under two circumstances: when a spectrum does not have a sufficient number of large peaks or when the nominal energy calibration is too far from the actual one. The former problem necessitates quality control of the internal energy calibration and a reversion mechanism. The latter usually indicates an incorrectly selected nominal calibration or an abrupt change in the channel-energy correspondence due to a hardware problem.

4.3.2 Shape Calibration

Information on the expected peak shape, especially the peak width, is needed in typical peak search algorithms, since the width of the search filter must reflect the peak width w. Due to the inherent properties of a Ge-spectrometer, w varies as a function of energy E, best described by the function $w(E) = \sqrt{a_0 + a_1 E + a_2 E^2}$ [22], where the non-negative parameters a_i depend on the quality of the spectrometer system. Since the channel-energy correspondence is very linear, the same functional form can be used when w is expressed in terms of channel C.

The shape information is provided by a shape calibration. If the calibration is used only in the peak search phase, the need for updates is not as obvious as when the peaks are also fitted on the basis of the shape calibration, so-called *fixed-width fitting* as opposed to *free-width fitting*. Nevertheless, peak search will also show degraded results when the real shape calibration drifts farther from the nominal one, i.e., the Type I and Type II error frequencies of peak search become higher than optimal.

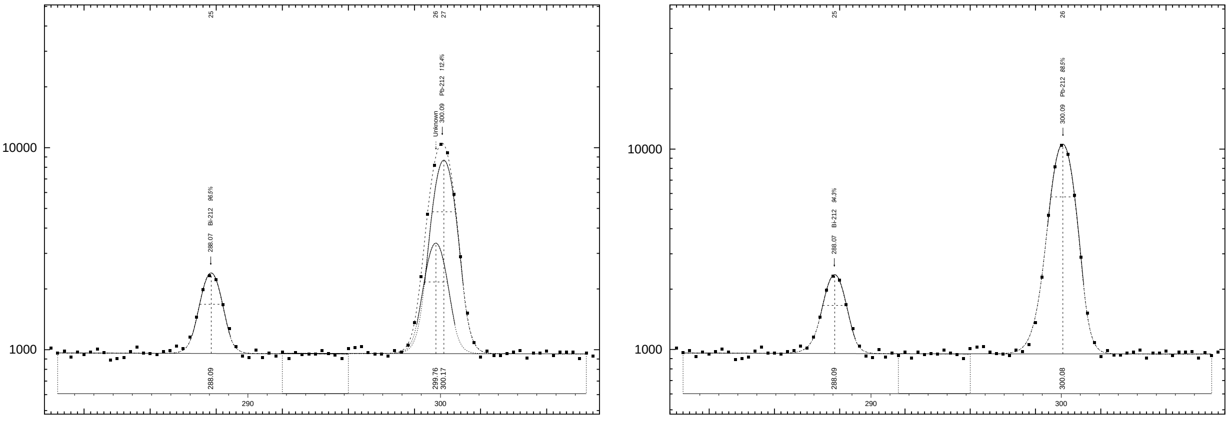

Figure 4.4: An example of peak analysis with an incorrect shape and a correct shape in pipeline analysis. On the left the nominal shape calibration is used, on the right the internal shape calibration.

Fig. 4.4 shows an example where a singlet peak is erroneously interpreted to be a doublet when using the nominal shape calibration. The smaller component of the doublet is actually added during peak fitting, not found by the search algorithm, but the case serves to illustrate the effects of an inaccurate shape calibration. An obvious solution to the problem would be to use free-width fitting in peak quantification. However, it introduces other problems, e.g., in resolving of close multiplets, that will be discussed in Sec. 6.1.

The drift of the peak shape should be quite modest in ideal laboratory conditions and it provides a basic quality control method: changes in peak width readily reveal hardware failures. When operating Ge-spectrometers in less ideal conditions, e.g., at stand-alone radiation monitoring stations or on mobile platforms, the shape of the peaks is more prone to change from one measurement to another. An internal shape calibration becomes essential for successful analysis. We apply a relatively simple calculation principle in the automated analysis pipeline of the Finnish National Data Centre (Sec. 2.2.4):

1. Search for the largest peaks in the spectrum using the nominal shape calibration.
2. Drop the annihilation peak and other peaks with anomalous shape parameters.
3. Fit the peaks keeping the shape parameters free.
4. Discard anomalously small or large parameters and those having large uncertainty estimates.
5. If the number of remaining parameters is sufficient, apply them as the shape calibration using either interpolation or suitable fitting functions. Otherwise, revert to the nominal calibration.

Failures in the internal shape calibration occur when the number of large singlet peaks is insufficient in the spectrum. According to test runs with a large set of air filter spectra, approximately 8 peaks with good-quality shape parameters suffices in pipeline analysis. If the number of good-quality peaks is lower, it is safer to apply the nominal shape calibration. It should be noted, however, that the limit has not been rigorously optimized.

4.3.3 Peak Efficiency Calibration

Peak efficiency calibration is not needed in peak analysis phases, but it is fundamental to the nuclide identification and activity determination phases. Efficiency differs from energy and shape calibrations in that an absolute efficiency calibration cannot easily be calculated internally, but it must be developed from calibration-source measurements, possibly accompanied with Monte Carlo simulations or other calculations. This means that the quality of the efficiency calibration given as input to the analysis has a direct effect on the quality of the nuclide identification results.

Modern spectrum analysis software packages usually implement efficiency calibrations with two alternative methods: a direct interpolation method or a fitted function method. A dozen of different functional forms have been postulated and implemented in analysis software packages for efficiency calibration, since the theoretically correct function is not known [49, p. 223]. In the Sampo program family, for example, the user can select between a log-log-linear efficiency interpolation and the following class of logarithmic polynomials that has proven to work decently:

$$\epsilon(E) = \exp\left(\sum_{i=0}^{N} a_i \ln(E)^i\right) ,$$

where ϵ is efficiency, E is energy, a_i are the fitted parameters and N is the degree of the polynomial. In a study with a relatively limited number of detectors, the optimum default value of N was empirically found to be 5 [90]. After operational experience with a larger batch of detectors, this choice has been confirmed in practice.

Since the class of efficiency calibration functions above has several free parameters, it is known to be relatively flexible, and therefore, its tails above and below the calibration points are known to be relatively “loose”. A small change in one of the data points may change the function parameters so that the function value at some 20 keV below the first data point may experience a change of hundreds percent. In order to prevent this kind of behavior, there should be a sufficient number of data points both above and below the maximum of the efficiency curve. In any case, efficiency extrapolation outside of the data range is prone to errors and should be avoided.

The linear extrapolation of data points is even more sensitive to the energy range of data points. When the efficiency values out of the data point interval are determined from the first two or last two data points, no changes in the curvature are accounted for, and thus, the extrapolated values diverge rapidly from the correct values if the curvature changes. This effect is especially strong at the low-energy end of the efficiency curve. Furthermore, any errors in the first two and the last two calibration points accumulate into the extrapolated values, making the divergence from the true value even more rapid. Of course, applying more than two points would add robustness in extrapolation.

In the efficiency calibration of Ge-spectrometers, it is a common practice to use one of the following energies as the lowest calibration point: 46 keV (^{210}Pb), 59 keV (^{241}Am) or 88 keV (^{109}Cd). All these energies are near the maximum of the efficiency curve, where the curvature changes rapidly, and thus, extrapolation of efficiencies to low energies is inaccurate. Especially when using modern broad-energy-range Ge-detectors, additional data points in the low-energy end are needed, but there are not so many alternatives. ^{125}I has a gamma-ray

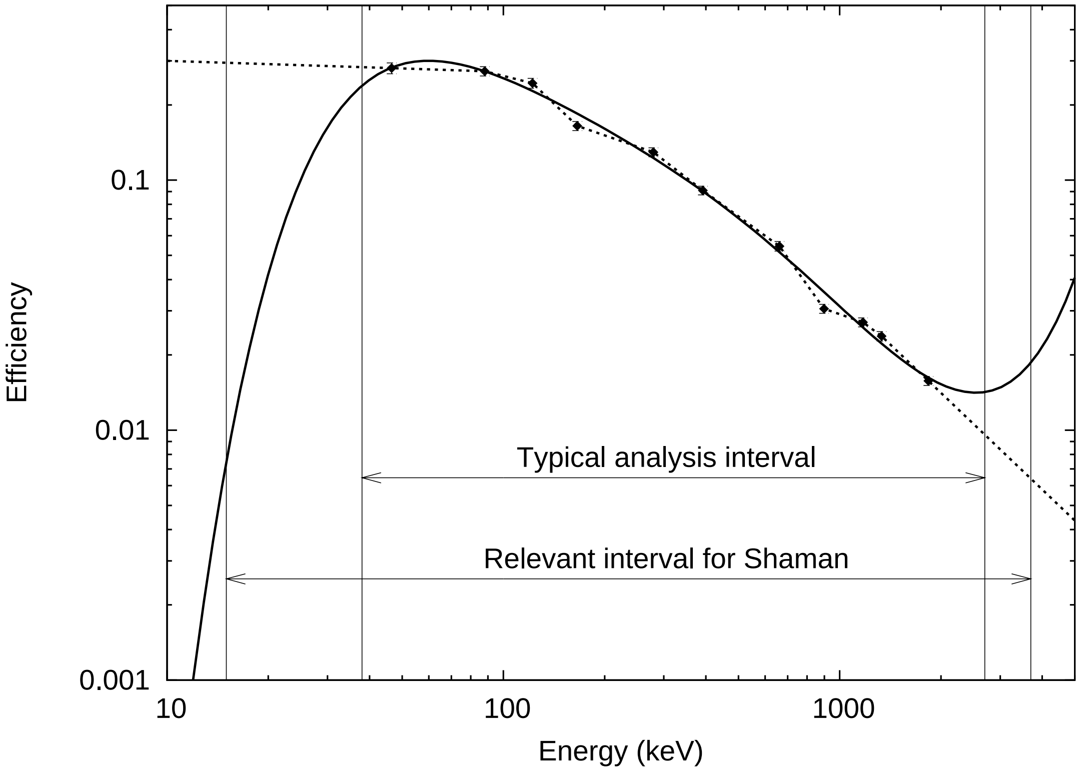

Figure 4.5: The efficiency vs. energy plot for an example measurement, illustrating the possible problems encountered in efficiency calibration. Efficiency data points are available for the energy interval [46, 1836] keV. A fifth-degree log-log-polynomial (solid line) has been fitted to the data points and interpolation of data points is shown with the dotted line. The correct efficiency values outside of the data interval are unknown, but the efficiency shall tend to zero at low and high energies. See discussion in text.

at 35 keV, but its half-life is only 59.4 d. ^{241}Am has relatively abundant X-ray peaks at 17–18 keV, but they are in a multiplet and their emission probabilities are quite inaccurate [49, p. 228]. Other nuclides with prominent X-ray peaks or Monte Carlo simulation may be a solution.

The linear log-log extrapolation method may lose its applicability some 10 keV below the lowest energy point as illustrated by the example of Fig. 4.5. The fitted efficiency function also fails occasionally in low energy extrapolation, but it tends to give more reasonable values to lower energies than linear extrapolation. The fitted function is therefore preferable.

The most common highest energy calibration point is the 1836 keV peak of ^{88}Y. This is usually sufficient for energies up to 3000 keV, because the log-log slope of the efficiency curve is almost constant between 200 and 3000 keV [49]. There is, however, a risk that the efficiency function fit, being a high-order polynomial, may exhibit excessive curvature above the last data point. Therefore, these fitted functions should not be used for quantitative purposes too far above the last data point, the limit typically varying from 100 to 1000 keV beyond this point. The example of this behavior shown in Fig. 4.5 is more dramatic than usual, but it has been chosen for illustration purposes. On the other hand, the linear log-

log extrapolation method usually works well above the data point interval, provided that the last two data points are sufficiently accurate. Therefore, it is the preferred method for efficiency extrapolation in the high-energy end. The accuracy of this method could be further improved, if a larger number of data points were used to define the slope.

If methods for coincidence summing correction and escape peak area prediction are applied in the nuclide identification task like in the expert system SHAMAN, reasonable efficiency estimates are required from some 15 keV in the low-energy end to about 1000 keV above the highest energy peak in the spectrum, regardless of possible cutoff energies applied in spectrum analysis. This is illustrated in Fig. 4.5. Extrapolated efficiencies are needed in coincidence summing calculations for all gamma-ray and X-ray energies present in the decay schemes of candidate nuclides. Escape peak area predictions, on the other hand, need to account for the possibility that the highest energy peak is a double escape peak, 1022 keV below a photopeak that is beyond the spectrum range. An efficiency estimate at the photopeak energy is required for a quantitative explanation of the escape peak. In SHAMAN's case, the fitted efficiency function is applied from the low energy end up to the highest data point, but linear log-log extrapolation is used above it.

The nuclides that are affected by these efficiency considerations are the ones that have prominent gamma-rays over a wide energy range. Common examples of such nuclides are ^{140}Ba, ^{208}Tl and ^{212}Bi. A typical consequence of an inaccurate efficiency calibration is that small peaks at either end of the energy range are left without an explanation, since their size seems not to be in the correct proportion to the more prominent peaks. Another consequence is that spurious nuclides are identified in order to explain the shares of peaks that are left unaccounted for by the correct nuclides. The quantitative results, i.e., nuclide activities and concentrations naturally reflect the errors in efficiency calibration even for single-line nuclides.

A non-trivial problem in efficiency calibration is coincidence summing, especially in close measuring geometries. Coincidence summing decreases the sizes of photopeaks corresponding to gamma-ray transitions in a cascade. Therefore, the efficiencies calculated on the basis of these peaks are underestimated. This happens with all calibration nuclides that are multiline emitters. Fig. 4.5 illustrates also this problem: the last four data points belong to two-line nuclides ^{60}Co (1173.2 and 1332.5 keV) and ^{88}Y (898.0 and 1836.1 keV) and the points are clearly below the log-log-linear line defined by the previous three calibration points of single-line emitters ^{137}Cs (661.7 keV), ^{113}Sn (391.7 keV) and ^{203}Hg (279.2 keV). This effect can be compensated for mathematically, but since the correction depends on the efficiency, the compensation method becomes iterative.

4.3.4 Total Efficiency Calibration

Total efficiency of a detector means the probability that a gamma-ray at a given energy interacts in the Ge-detector in such a way that a pulse is recorded anywhere in the spectrum. In peak efficiency, only the pulses recorded in the full-energy peak are taken into account and therefore, total efficiency is always larger than or equal to peak efficiency. Peak efficiency is always needed for converting peak areas to emission rates and activities, but the total efficiency calibration is needed only in coincidence summing calculations.

Total efficiency calibration can be made empirically. However, this requires a number of single-line emitters in the source geometry in use. Some multiline emitters like ^{60}Co can be used as pseudo-single-line emitters and the two-line emitter ^{88}Y is commonly used to provide a data point at 1836.1 keV. This is done by estimating the total efficiency at 898.0 keV from other data points and subtracting its contribution from the spectrum counts [49].

One complication when calculating the total number of counts in a spectrum is the effect of low-level discriminator (LLD) that suppresses noise pulses from the spectrum. The LLD makes the contents of the first N channels zero and thus, the LLD value would have an effect on the total efficiency. In order to avoid this ambiguity, the channel contents in the LLD region are customarily estimated from those above the cut-off [49]. The most straightforward method is to calculate the average number of counts in 5–20 channels just above LLD and assume this number of counts for each channel below LLD.

An empirical total efficiency calibration requires quite an effort, and especially if there are several source geometries in use, it can become impractical. There are three basic alternatives for an empirical total efficiency calibration:

1. estimating total efficiency from peak efficiency utilizing a suitable parameterization,
2. analytical calculation of total efficiency from basic physical principles and detector dimensions,
3. calculating total efficiency with Monte Carlo simulation.

The effort needed by these methods increases in the given order but so does the obtainable accuracy. The last alternative is not discussed here, because it practically requires modeling of the source-detector setup, after which the total efficiency is obtained with a straightforward calculation.

Estimation of Total Efficiency from Peak Efficiency

We have developed a parameterized approximation of total efficiency based on the peak efficiency calibration and detector size and type. Parameters have been obtained from semi-empirical peak and total efficiencies for 10 different detectors of STUK (Finnish Radiation and Nuclear Safety Authority) used in a standardized air filter measurement setup [91]. These calibrations are based on measurements, but the data points are given on energies of even tens or hundreds keV's by interpolating the measured data points. This makes calculation of total-to-peak-efficiency ratios at a sufficient number of energies straightforward. The ratios were plotted as a function of energy in Fig. 4.6 and the following conclusions were boldly drawn, despite the quite limited amount of data:

1. Relative efficiency of a detector is an essential parameter: all $\epsilon_r \approx 30\,\%$ detectors behave approximately similarly and so do all $\epsilon_r \approx 40\,\%$ detectors.
2. The curves of n-type detectors seem to behave like those of p-type detectors, but with $\epsilon_r \sim 10\,\%$ larger.

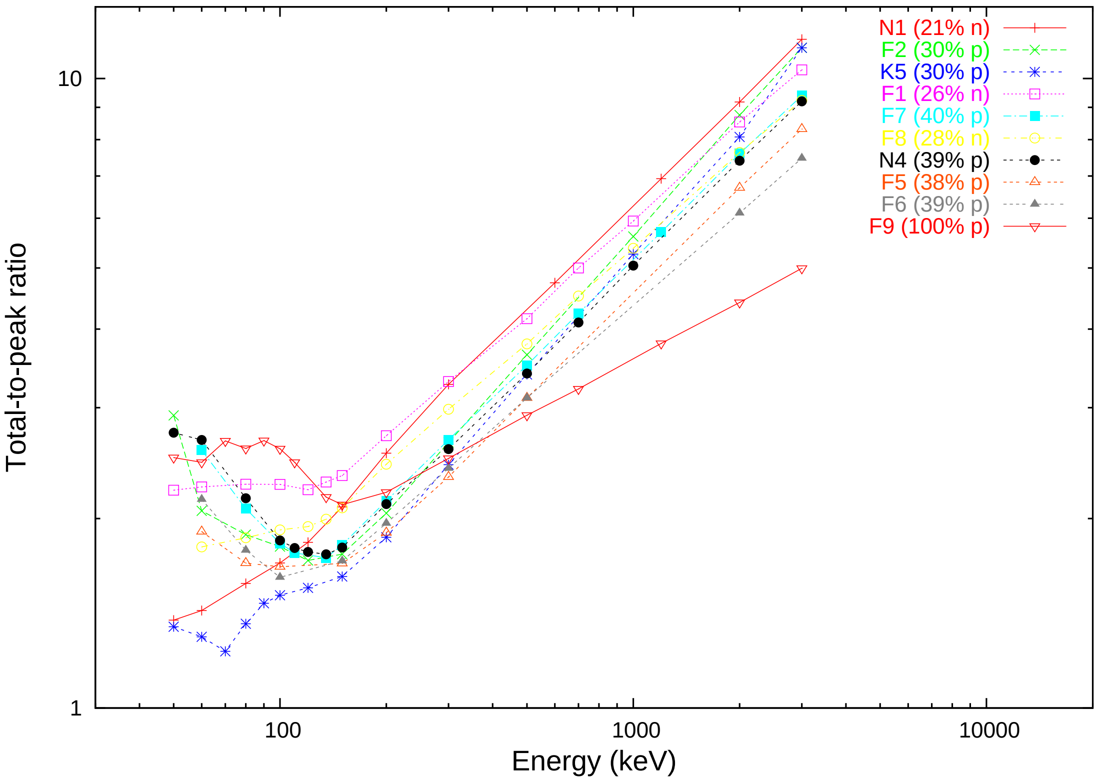

Figure 4.6: Total-to-peak-efficiency ratio as a function of energy for 10 different detectors. The ratio has been calculated from semi-empirical efficiency calibrations in a standardized air filter measurement setup.

3. The behavior in the low-energy end looks quite random, so it should be smoothed. A constant value below 100 keV is the safest choice.

The uppermost 4 curves were pooled together to obtain an average curve for $\epsilon_r = 30\,\%$ detectors and the next 5 curves to obtain an average for $\epsilon_r = 40\,\%$ detectors. The curve of detector F9 is used for all $\epsilon_r = 100\,\%$ detectors. The resulting curves are shown in Fig. 4.7. These curves are interpolated for detectors with $\epsilon_r \in [30\,\%, 100\,\%]$, but extrapolation is not applied. Instead, the 30 % curve is used for relative efficiencies below 30 % and the 100 % curve for relative efficiencies above 100 %.

This approximation usually gives total efficiencies with sufficient accuracy for coincidence correction calculations. This is partly due to the fact that coincidence correction calculation is relatively insensitive to the accuracy of total efficiency. It is only included in terms of the form $(1 - \epsilon_t)$, so even if ϵ_t was of the order of 0.2 and it had a 20 % uncertainty, the uncertainty of $(1 - \epsilon_t)$ would be only 5 %. Nevertheless, an empirical or calculated total efficiency calibration is always recommended if available. Especially, if an n-type detector or a p-type detector with enhanced energy range is used, the efficiency at energies below 100 keV is so high that its accuracy has a profound effect on the coincidence correction calculations. This estimation method for total efficiency does not seem to work very well in these cases.

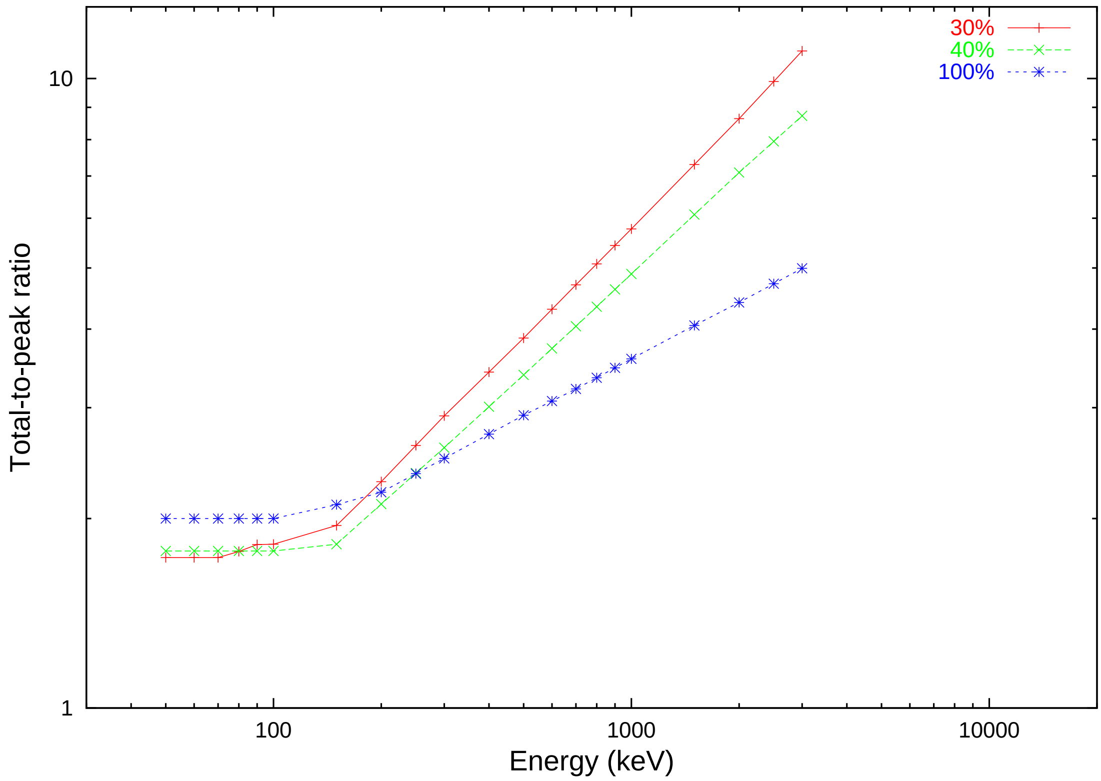

Figure 4.7: Parameterized total-to-peak-efficiency ratio as a function of energy, using the relative efficiency ϵ_r of the detector as the essential parameter.

Calculation of Total Efficiency with Numerical Integration

A method for calculating total efficiency from basic physical principles and detector dimensions has been presented in Ref. [92] for point sources and true coaxial Ge-detectors. Since the probability P for gamma-ray interaction in the detector is $P = 1 - \exp(-\mu S)$ where μ is the total absorption coefficient and S is the slant thickness of (path length in) the detector, total efficiency can be expressed as:

$$\epsilon_t = \frac{1}{4\pi} \int [1 - \exp(-\mu S)] \; d\Omega \quad , \tag{4.1}$$

where $d\Omega$ is an infinitesimal solid angle over which S may be considered constant. Ref. [92] derives the calculation for a true coaxial detector where the inactive core of the detector extends through the detector, but the derivation can be extended for a closed-end detector with a core extending to any distance from the back face. The following formula is obtained:

$$\begin{aligned}
\epsilon_{t0} &= \frac{D}{2} \Bigg\{ \int_0^{R_2} \frac{r}{(r^2 + D^2)^{3/2}} \, dr \\
&+ \int_{R_1 D/(D+t)}^{R_1 D/(D+t-u)} \frac{r}{(r^2 + D^2)^{3/2}} \exp\left[\frac{-\mu(R_1 - r)(r^2 + D^2)^{1/2}}{r}\right] dr \\
&- \int_0^{R_1 D/(D+t-u)} \frac{r}{(r^2 + D^2)^{3/2}} \exp\left[\frac{-\mu(t - u)(r^2 + D^2)^{1/2}}{d}\right] dr
\end{aligned}$$

$$- \int_{R_1 D/(D+t)}^{R_2 D/(D+t)} \frac{r}{(r^2+D^2)^{3/2}} \exp\left[\frac{-\mu t (r^2+D^2)^{1/2}}{d}\right] dr$$

$$\left. - \int_{R_2 D/(D+t)}^{R_2} \frac{r}{(r^2+D^2)^{3/2}} \exp\left[\frac{-\mu (R_2 - r)(r^2+D^2)^{1/2}}{r}\right] dr \right\} \quad , \tag{4.2}$$

where R_1 is the radius of the core, R_2 the radius of the detector, t is the length of the detector, u is the length of the core and D is the source-detector distance. The first term on the right-hand side can be integrated in closed form $(1/D - 1/(R_2^2 + D^2)^{1/2})$, but the rest must be evaluated numerically.

It can be seen from the plots of Ref. [92] that the total efficiency curves produced by the method are not realistic in the low-energy end: the curves have a constant value below ~100 keV. This indicates the fact that attenuation in the detector housing and the inactive layer of the detector are not accounted for. This attenuation is a separate phenomenon that can be incorporated into the calculation method as an additional factor. Since the attenuating layers are relatively thin, the following simple correction suffices:

$$\epsilon_t = \epsilon_{t0} \times \exp(-\mu_{p,w} t_w) \exp(-\mu_p t_i) \quad , \tag{4.3}$$

where $\mu_{p,w}$ is the photoelectric absorption coefficient of the window material (typically Al or Be) and t_w its thickness, and μ_p is the photoelectric absorption coefficient of germanium and t_i the thickness of the inactive Ge-layer.

The example of Fig. 4.8 illustrates the accuracy of this phenomenological approximation. We acquired the empirical data with a p-type Ge-detector with $\epsilon_r = 16.5\,\%$ using a number of point-like calibration sources at a distance of 100 mm from the detector end-cap. Empirical peak efficiencies were determined with the SAMPO program and the total efficiencies of single-line emitters on a spreadsheet. The effects of the low-level discriminator and non-negligible background count rate were compensated for.

When using the nominal detector dimensions, the calculated total efficiencies are within 10 % of the empirical values in the high-energy region, but the efficiency estimate at 60 keV is twice the empirical value. However, the dimensions of older Ge-detectors are usually quite imprecisely known, so it is well motivated to adjust the nominal values to obtain a better correspondence with empirical values. The high-energy values were made to match the empirical values by increasing the source-detector distance with 6 mm. The 60 keV point was reproduced when the inactive Ge-layer was increased from 0.7 mm to 1.4 mm. With these two adjustments, the calculated ϵ_t curve was made to go through the empirical points.

It is interesting to note that when the same detector was modeled with Monte Carlo simulation in Ref. [93], similar adjustments of dimensions were found necessary. This confirms the doubt that the physical dimensions presented in the manufacturer's specifications are quite inaccurate, or at least they do not account for the inhomogeneous electric field of the detector, leading to nonuniform charge collection. This would also suggest that the numerical integration method for total efficiency determination could match in accuracy with Monte Carlo simulation, which is substantially more computing intensive, but more evidence is required to make a proof.

Fig. 4.8 also presents a total efficiency curve calculated with the parameterized total-to-peak ratio method presented earlier. It can be seen that the empirical points at > 500 keV are

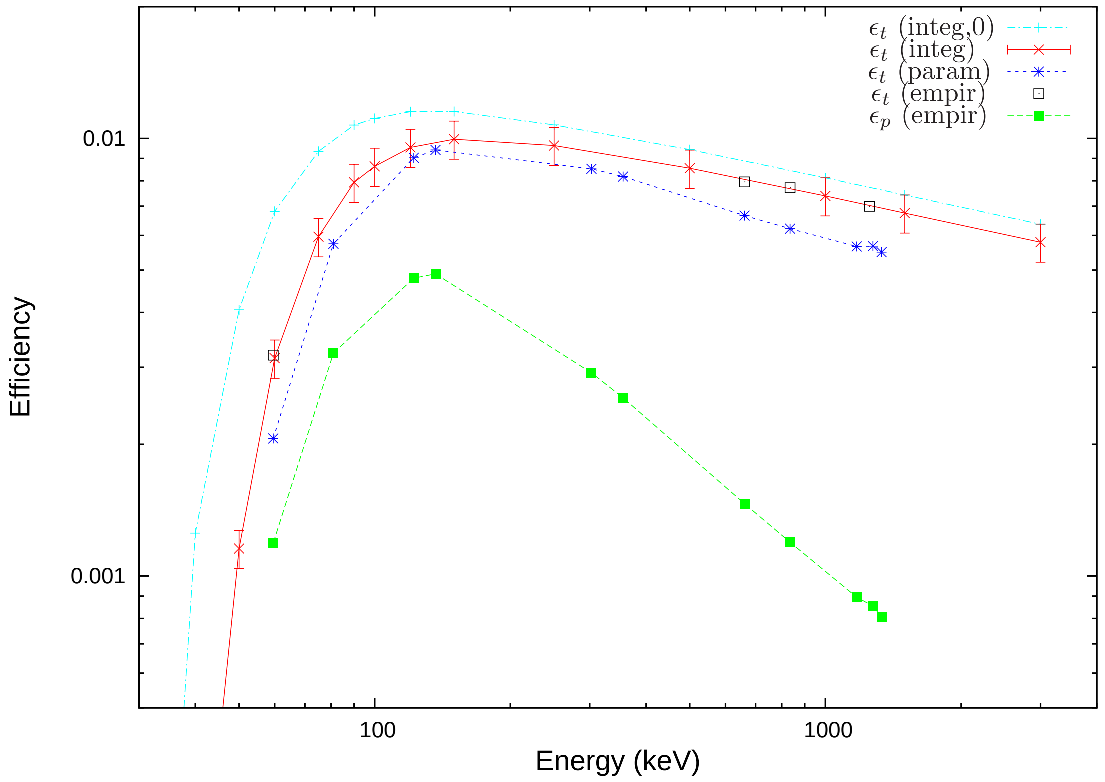

Figure 4.8: Calculated total efficiency curves vs. empirically determined values. The curve labeled "ϵ_t (integ,0)" was obtained with the numerical integration method using nominal detector dimensions and the curve labeled "ϵ_t (integ)" with slightly modified dimensions. The curve labeled "ϵ_t (param)" was obtained with the parameterized total-to-peak ratio. The points labeled "ϵ_t (empir)" are empirically determined total efficiencies and the points labeled "ϵ_p (empir)" empirically determined peak efficiencies.

$\sim$20 % above and the 60 keV point $\sim$50 % above the estimated curve. A discrepancy in this direction was to be expected, because the curve for detectors with $\epsilon_r = 30\,\%$ had to be used for this $\epsilon_r = 16.5\,\%$ detector. Nevertheless, even this total efficiency calibration would be applicable to computing a first approximation of the coincidence correction factors.

4.4 Decision, Detection and Determination Limits

4.4.1 Calculation of the Limits

When estimating the limits of detecting a radionuclide signal, a classic paper by Currie, Ref. [94], is practically always quoted. It defines the basic quantities that are to be applied in radiation measurements. Although the original paper is a purely statistical approach that is directly applicable for single-channel analyzers (SCA), many researchers, including us, apply it to pulse-height spectra measured with a multi-channel analyzer (MCA). This requires an SCA-MCA analogy to be defined. It should be noted that these analogies are always approximations, but more importantly, they do not account for the actual methods used in MCA spectrum analysis, i.e., peak search and peak area determination. Thus, the calculated limits are not to be interpreted too strictly in MCA spectrum analysis: the actual limit of detection most probably differs from the value given by the SCA analogy.

When judging the signal observation possibilities in nuclear measurements, three characteristic quantities can be defined [94]:

- *Decision limit* L_C guides a decision whether there are excess counts registered among the gross number of counts in a region-of-interest (ROI).

  The decision limit L_C is defined as such a value that if the observed net signal $N_n > L_C$, it is concluded that there exist a non-zero number of excess counts in the ROI. Otherwise, it is concluded that no net counts exist. If in reality only blank counts exist in the ROI, this decision rule leads to the incorrect conclusion that net counts exist (error of the first kind) with the probability α selected in advance.

- *Detection limit* L_D tells the minimum expectation value of excess counts in a ROI which can be proven in this kind of a measurement. This value guides a decision whether the measurement method fulfills the requirements defined in advance.

  The detection limit L_D is the smallest possible expectation value of the net signal, for which the L_C presented above makes an error of the second kind, i.e., claiming no net signal although there is some, at the maximum probability β selected in advance. The detection limit is to be used to prove whether a measuring setup works for a certain purpose. This is done by comparison of the detection limit with a predefined guide value.

- *Determination limit* L_Q tells the minimum expectation value of excess counts in a ROI which can be proven in this kind of a measurement quantitatively, i.e., at a relative accuracy requested in advance.

  The determination limit L_Q is the smallest possible expectation value of the net signal, for which the quantitative result is satisfactorily close to the true value. The maximum acceptable relative standard deviation, denoted as $1/k_Q$, is to be selected in advance.

The following equations can be derived for the three limits of detection when the blank measurement has an uncertainty comparable to that of the gross measurement [94]:

$$L_C = k_\alpha \sqrt{2}\sigma_B \ , \tag{4.4}$$

$$L_D = L_C + \frac{k_\beta^2}{2} + k_\beta \sqrt{\frac{k_\beta^2}{4} + L_C + \frac{L_C^2}{k_\alpha^2}} \ , \tag{4.5}$$

$$L_Q = \frac{k_Q^2}{2} + k_Q \sqrt{\frac{k_Q^2}{4} + 2\sigma_B^2} = \frac{k_Q^2}{2} + k_Q \sqrt{\frac{k_Q^2}{4} + \frac{L_C^2}{k_\alpha^2}} \ , \tag{4.6}$$

where σ_B is the standard deviation of the blank count B. Customarily the risk levels for the error of the first and second kind are chosen equal or $\alpha = \beta$, implicating $k_\alpha = k_\beta = k$. This simplifies the L_D equation substantially:

$$k_\alpha = k_\beta = k \quad \Rightarrow \quad L_D = k^2 + 2L_C \ .$$

It can also be seen that for all but the smallest numbers of counts $L_Q \approx \frac{k_Q}{k_\alpha} L_C$.

4.4.2 Calculation of the Limits in SHAMAN

The equations above are rigorously valid only for an SCA. When applying them for MCA measurements, one needs to define an analogy between the SCA window and an ROI of MCA-spectrum where the Gaussian peak has the width w. The optimal length of the ROI for detection limit calculation can be derived by maximizing the signal-to-noise ratio. This has been done in Ref. [5] for a singlet peak on a constant baseline. The result of the optimization is that the ROI should extend $1.4w$ to both directions from the peak centroid.

Applying the optimum width of ROI when the average number of counts per channel is b, the blank area estimate becomes $B = 2.8bw$. According to Poisson statistics, the corresponding standard deviation is

$$\sigma_B = \sqrt{2.8bw} \ . \tag{4.7}$$

It should further be noted that the optimum width of ROI includes only $F = 83.8\,\%$ of the Gaussian peak area. If the peak area is obtained by integrating over $(-\infty, +\infty)$, it shall therefore be compared to the compensated decision limit

$$L_C(\text{SHAMAN}) = k_\alpha \sqrt{2 \times 2.8bw}/F = 4.66\sqrt{bw} \ , \tag{4.8}$$

where the value $k_\alpha = 1.65$, corresponding to the risk level $\alpha = 0.05$ for a Gaussian distribution,[b] has been applied in the last phase. Naturally, the need for compensation and its calculation depends on the peak area calculation method. When applying this compensated L_C, L_D becomes automatically compensated as it depends on L_C, but for L_Q the compensation needs to be done explicitly unless applying the latter form of Eq. 4.6.

As was already stated, the L_C calculation method for MCA measurements attempts to emulate the method for an SCA, but it is only an approximation. This has led to many discussions on the constants involved in literature. If one accepts the approximate nature of the equations above, then $L_C = C\sqrt{bw}$ where the constant C is to be chosen either by investigating the actual decision limit of the peak analysis method with artificial spectra or by making an administrative decision. An example of the latter is the CTBTO decision to use the calculation interval 2.5×FWHM or $5.89 \times w$ as the summing interval and $k_\alpha = 1.65$. The wide interval makes $F = 1.0$ and the equation for L_C becomes

$$L_C(\text{CTBTO}) = k_\alpha \sqrt{2 \times 5.89bw}/F = 5.66\sqrt{bw} = 1.22 L_C(\text{SHAMAN}) \ .$$

4.4.3 Comparison of Calculation Methods

Alternative calculation methods for limits of detection have been presented in literature. Those presented in Refs. [95], [96] and [97] will be investigated in the following in parallel with SHAMAN's "compensated Currie-analogy" presented above. Only the decision limit L_C will be discussed, because L_D and L_Q are dependent on it.

Let us apply the notation above and simplify the comparison by making $k_\alpha = k_\beta = k$ and assuming a constant baseline. Furthermore, let us assume that the blank signal is determined

---

[b] With small peak areas, the Poisson distribution shall be applied instead of Gaussian.

from an equally wide spectrum interval as the gross signal in equations of Refs. [96] and [97]. The following equations for decision limit are to be compared:

$$
\begin{aligned}
L_C(\textsc{Shaman}) &= k\sqrt{2B}/F = k\sqrt{2 \times 2.8 \times bw}/F && (4.9)\\
L_C(\text{Aničin\&Yap}) &= \frac{k^2}{2}\left(1+\sqrt{1+\frac{8B}{k^2}}\right) = \frac{k^2}{2} + k\sqrt{\frac{k^2}{4} + 2 \times 2.55 \times 2.35 \times bw} && (4.10)\\
L_C(\text{DIN-25482-2}) &= \frac{k^2}{2} + k\sqrt{\frac{k^2}{4} + 2B} = \frac{k^2}{2} + k\sqrt{\frac{k^2}{4} + 2 \times 2.5 \times 2.35 \times bw} && (4.11)\\
L_C(\text{DIN-25482-5-e}) &= \frac{k^2}{2}\left(1+\sqrt{1+\frac{8B}{k^2}}\right) = \frac{k^2}{2} + k\sqrt{\frac{k^2}{4} + 2 \times 2.5 \times 2.35 \times bw} && (4.12)\\
L_C(\text{DIN-25482-5-a}) &= k\sqrt{2B} = k\sqrt{2 \times 2.5 \times 2.35 \times bw} && (4.13)
\end{aligned}
$$

It can be seen that the forms of L_C(Aničin&Yap) [95], L_C(DIN-25482-2) [96] and L_C(DIN-25482-5-e) [97] actually correspond to that of L_Q of Eq. 4.6. On the other hand, the approximate version L_C(DIN-25482-5-a) [97] corresponds to L_C of Eq. 4.4. L_C(DIN-25482-2) and L_C(DIN-25482-5-e) have converged identical with this selection of parameters, but the basic equations are in fact different.

All L_C-versions above are seen to tend toward $C\sqrt{bw}$ when $bw \gg k$. However, the factor C varies in the following way when using the risk level $\alpha = 0.025$ that is recommended in Ref. [97] and that corresponds to $k = 1.96$ in a Gaussian distribution: $C =$ 5.53 / 6.79 / 6.73 / 6.73 / 6.73 from top to bottom. This is due to different recommendations for the blank summing interval and the factor F in the first equation.

However, even if the asymptotic behavior of the L_C-equations is the same and the peak width parameter w can be considered a fixed constant given from outside, a large discrepancy is found to be caused by different calculation methods for the average baseline b in the ROI. The recommended methods for baseline determination are the following:

1. L_C(Shaman) [5]: Use the fitted baseline and integrate it over $2.8w$ at the location of the peak. If the baseline has not been fitted, it is approximated by summation of channel contents over exactly $2.8w$.

2. L_C(Aničin&Yap) [95]: Determine the baseline as an average of the two channels adjacent to the ROI whose width is 2.55× FWHM.

3. L_C(DIN-25482-2) [96]: Determine the baseline from a symmetrical region adjacent to the ROI whose width is 2.5× FWHM. A statistical test is provided in the document for determining the baseline region as wide as possible.

4. L_C(DIN-25482-5-e) [97]: Determine the baseline from a symmetrical region adjacent to the ROI whose width is 2.5× FWHM. The total width of the baseline region should be between one and ten times the ROI width, but in any case free of any tails of neighboring peaks.

   An additional note in Ref. [97] states, however, that the applicability of the presented calculation formulas is widest when the baseline region width is chosen equal to the ROI width. This can be considered as a recommendation.

5. L_C(DIN-25482-5-a) [97]: Same baseline as for the previous one.

   Since this approximate method consistently underestimates the “exact value” or L_C(DIN-25482-5-e) with a few percent, there is no point using it instead of the exact alternative with current computing capacity.

We wrote a computer program to calculate the decision limits by each of these methods and applied it on a set of synthesized spectra with known contents and correct statistics. An example of the results is shown in Fig. 4.9 that presents a representative region of 200 channels with three small peaks. In this spectrum, the baseline is 100 counts, the peaks 150 counts each and at 80 channel intervals. The peak width parameter is $w = 2.0$ channels, corresponding to FWHM = 4.71 channels. The constant values in Fig. 4.9 shown with dashed lines would be obtained with each method for a spectrum with exactly 100 counts on each channel.

The following conclusions can be drawn from this figure:

1. L_C(SHAMAN) [5]: The channel-to-channel fluctuations are relatively smooth and the standard deviation of the L_C-values is 3.2 %. The values are close to the “ideal” value except in the proximity of peaks. If these peaks were fitted, the decision limit would be calculated from the fitted baseline function leading to a smooth behavior also in these regions.

2. L_C(Aničin&Yap) [95]: The channel-to-channel fluctuations are relatively large, although the standard deviation of the L_C-values is only 4.0 %. This is because the baseline is calculated as an average of only two channels. The peaks are manifested by two maxima around the centroid channel, which is in contrast to L_C(SHAMAN).

3. L_C(DIN-25482-2) [96]: The channel-to-channel fluctuations are very smooth in the areas where a wide baseline summing interval is utilized. However, the selection rule for the interval is unstable: it experiences large changes especially close to spectrum peaks but also elsewhere and this leads to channel-to-channel fluctuations of up to 100 %. The standard deviation of the L_C-values is 23 %. Thus, the applicability of this method can be questioned, at least in automated analysis of high-resolution gamma-ray spectra.

4. L_C(DIN-25482-5-e) [97]: The channel-to-channel fluctuations are relatively smooth and the standard deviation of the L_C-values is 2.2 %. The values are close to the “ideal” value except in the proximity of peaks. The peaks are manifested by two maxima around the centroid channel, which is in contrast to L_C(SHAMAN). The behavior is very similar to L_C(Aničin&Yap), but the wider baseline summing interval stabilizes L_C(DIN-25482-5-e) and makes it a preferable method of these two.

These considerations practically leave two alternative methods for decision limit calculation in gamma-ray spectrometry, namely L_C(SHAMAN) and L_C(DIN-25482-5-e). Their main differences are two: first, the wider summing interval of the DIN-method leads to a constant difference by a factor of ∼1.25, and second, the behavior of L_C in the proximity of spectral peaks or bumps is different. The latter is due to the selection of summing interval:

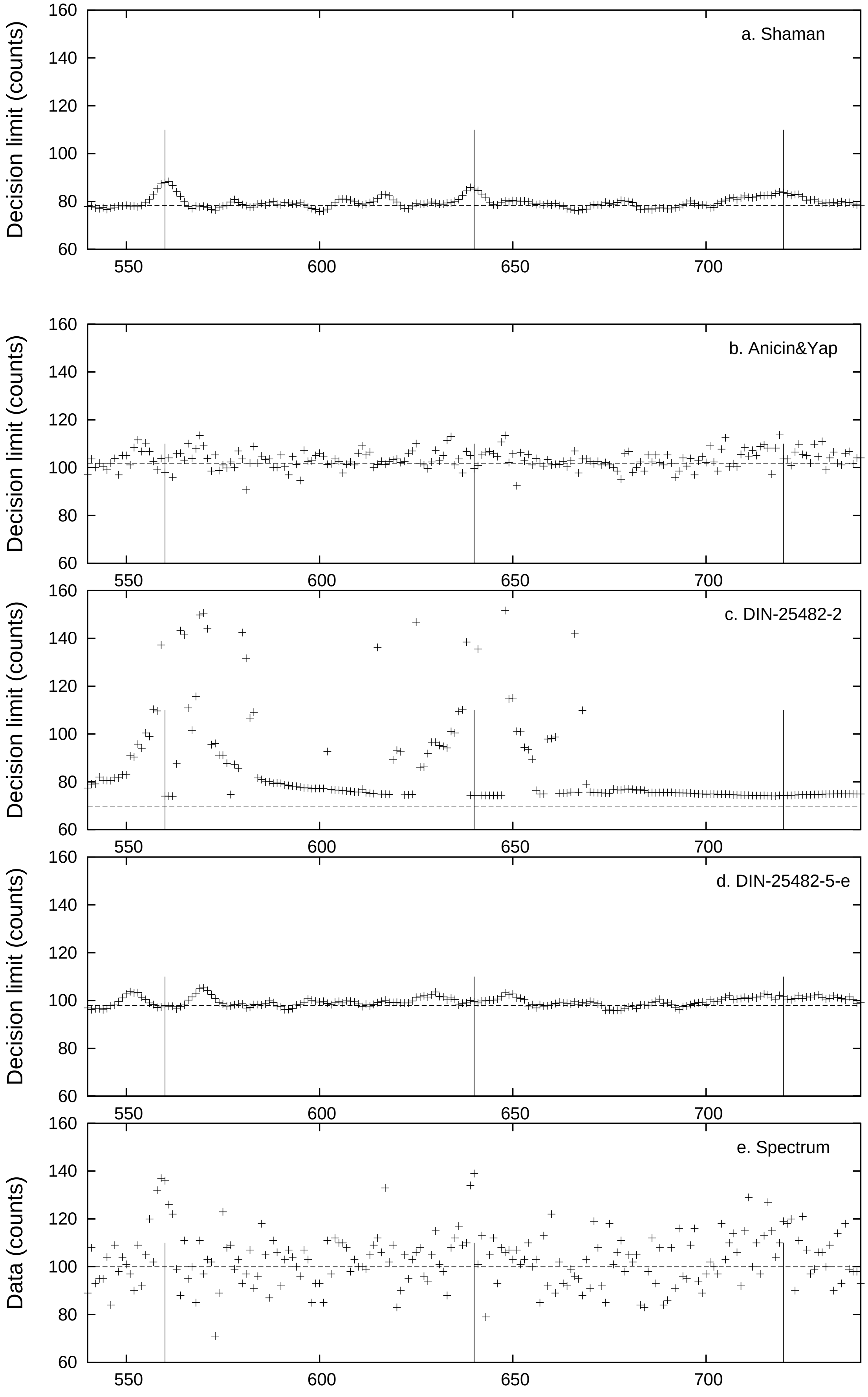

Figure 4.9: Decision limits calculated with different methods a.–d. for the synthesized spectrum in e. with small peaks at channels 560, 640 and 720 on a constant baseline of 100 counts. The dashed lines show the constant result for a spectrum with 100 counts on each channel.

either under the assumed peak like in the SHAMAN method or on its both sides like in the DIN-method. Both selections can be motivated, so the choice of method can be based on preference or regulations.

The factor C has been interpreted here more or less as a free scaling factor. It should be explicitly evaluated how well the two recommended methods obey the nominal risk level. This can be accomplished with synthesized spectra with constant baselines and known peak contents. In the calculations above both methods were applied with a nominal risk level of $\alpha = 0.025$. Obviously, one method must be closer to it than the other.

4.5 Annihilation Escape Peaks

Radionuclide identification is based on a comparison of photopeak energies and areas to the energies and emission probabilities in a reference library. Other types of peaks present in a typical gamma-ray spectrum complicate the identification task somewhat. Methods for identification and quantification of annihilation escape peaks, random sum peaks, coincidence sum peaks and X-ray escape peaks are presented in Secs. 4.5–4.8.

4.5.1 Calculation Method

Heydorn and Rhee [98] have presented a method for advance prediction of escape peak areas. The method is based on the theory of pair production and measurement of several spectra of nuclides with prominent gamma-ray energies in the range 1000–4000 keV at different distances. Heydorn and Rhee concluded that the distance does not effect the ratios of escape and photopeak areas. Thus, the single-escape-to-photopeak ratio R_1 and the double-escape-to-photopeak ratio R_2 of a Ge detector are functions of gamma-ray energy only. It is also known that the single-to-double-escape-peak ratio R' is a detector-dependent constant [22, 49], and hence, the energy dependences of R_1 and R_2 are of the same functional form.

Heydorn and Rhee state that the functional form of $R_1(E_\gamma)$ and $R_2(E_\gamma)$ can be deduced from the shape of the pair production cross section κ. The following rapidly convergent series expansion for κ as a function of energy $k = E_\gamma/511.034\,\mathrm{keV}$ has been derived by Maximon [99]:

$$\kappa \propto \left(\frac{k-2}{k}\right)^3 \left(1 + \frac{1}{2}\rho + \frac{23}{40}\rho^2 + \frac{11}{60}\rho^3 + \frac{29}{960}\rho^4 + \ldots\right) , \tag{4.14}$$

where

$$\rho = \frac{2k-4}{2+k+2\sqrt{2k}} . \tag{4.15}$$

When using the terms shown explicitly in Eq. 4.14, the relative error at $k = 4$ is $-0.011\,\%$, and the error becomes even smaller at lower energies. Another expansion has been derived by Maximon for energies above $k = 4$, but the expansion above suffices well for our purposes as it gives values within 0.5 % of the correct value for energies up to 5000 keV.

Spectrum	Measuring time (h)	Counting rate (cps) (dead time)	Shaping time (μs)	Source-detector distance (mm)	Nuclide contents
`esc1`	4.0	2 420 (6.6 %)	6	150 ± 2	^{22}Na (4.57 kBq), ^{60}Co (43.2 kBq), ^{152}Eu (297 kBq), ^{208}Tl (0.324 cps)
`esc2`	58.7	2 420 (6.5 %)	6	150 ± 2	^{22}Na (4.57 kBq), ^{60}Co (43.2 kBq), ^{152}Eu (297 kBq), ^{208}Tl (0.324 cps)
`sum1`	16.2	1 660 (4.5 %)	6	150 ± 2	^{60}Co (43.2 kBq), ^{137}Cs (281 kBq), ^{208}Tl (0.324 cps), ^{241}Am (381 kBq)
`sum2`	5.0	6 230 (16.3 %)	6	60 ± 2	^{60}Co (43.2 kBq), ^{137}Cs (281 kBq), ^{208}Tl (0.324 cps), ^{241}Am (381 kBq)
`sum3`	5.0	6 270 (6.2 %)	2	60 ± 2	^{60}Co (43.2 kBq), ^{137}Cs (281 kBq), ^{208}Tl (0.324 cps), ^{241}Am (381 kBq)
`meas14`	0.17	11 480 (33.1 %)	6	100 ± 2	^{116m}In (600 kBq)
`stuk6`	1.5	11 710 (? %)	6	close	^{85m}Kr (1.4 kBq), ^{88}Kr (0.54 kBq), ^{88}Rb (0.63 kBq), ^{133}Xe (42 kBq), ^{133m}Xe (2.5 kBq), ^{135}Xe (60 kBq)

Table 4.1: The basic data of the test spectra for escape and random sum peak studies.

According to Ref. [98], a straight line is achieved when the double-escape-to-photopeak ratio R_2 multiplied by the corresponding photopeak efficiency is plotted as a function of κ. This line should go through the origin since the double escape peak vanishes at $\kappa = 0$ by definition. The slope s of this line is the only detector characteristic needed to predict R_2 at all energies. It should be noted that R_2 is a detector-specific constant but dependent on energy, whereas s is independent of energy but dependent on the measuring geometry.

When the single-to-double-escape-peak ratio R' is known, R_1 can be predicted from R_2 or vice versa. According to Ref. [100], the value of R' can be determined from the active volume of the detector using a linear empirical relationship. Thus, only one well-defined escape peak is sufficient for applying the prediction method, but for a better accuracy, a source with at least two high energy gamma-rays (e.g., ^{24}Na with energies 1369 and 2754 keV) is recommended.

4.5.2 Experimental Results

We evaluated the method for advance prediction of escape peak areas experimentally with the help of five gamma-ray spectra named `esc1`, `esc2`, `sum1`, `sum2`, and `sum3` [V]. They were measured using some common calibration nuclides and an EG&G Ortec Poptop HPGe-detector (p-type, $\epsilon_r = 16.5\,\%$, FWHM(122 keV) = 0.7 keV, FWHM(1332 keV) = 1.7 keV). The basic data of the test spectra are presented in Table 4.1 that also includes data for two other spectra used for evaluation of random summing, named `meas14` and `stuk6`. The spectrum `meas14` was measured with the above spectrometer of an irradiated indium sample and the spectrum `stuk6` is a spectrum of a gaseous reactor sample obtained from STUK (Finnish Radiation and Nuclear Safety Authority).

In Table 4.1, measuring times are the approximate live times and counting rates the total counting rates of the spectra. Dead time percentages are shown below counting rates, with

Spectrum	E_p (keV)	E_s (keV)	E_d (keV)	R_1	R_2	R'
`esc1`	1299.1		277.1		0.02845 (32.8 %)	
	1332.5	820.9	310.5	0.00245 (19.8 %)	0.00365 (20.0 %)	0.6706 (28.1 %)
	1408.1	897.0	386.0	0.00414 (11.3 %)	0.00892 (5.5 %)	0.4635 (12.5 %)
	2614.3	2103.4	1592.6	0.16081 (5.1 %)	0.15190 (7.4 %)	1.0586 (8.9 %)
`esc2`	1299.0		277.2		0.03122 (6.7 %)	
	1332.3	821.3	310.5	0.00340 (7.7 %)	0.00381 (5.6 %)	0.8908 (9.5 %)
	1407.9	896.9	386.0	0.00554 (4.7 %)	0.00792 (2.9 %)	0.6987 (5.5 %)
	2614.1	2103.1	1592.3	0.15325 (1.5 %)	0.14977 (2.2 %)	1.0233 (2.5 %)
`sum1`	1332.2	821.2	310.5	0.00451 (7.2 %)	0.00532 (9.0 %)	0.8476 (11.5 %)
	2614.0	2103.1	1592.2	0.14564 (2.5 %)	0.16262 (2.8 %)	0.8956 (3.7 %)
`sum2`	1333.0	821.9	310.6	0.00323 (8.6 %)	0.00519 (7.6 %)	0.6216 (11.5 %)
	2615.2	2104.1	1593.1	0.13958 (6.8 %)	0.15443 (12.1 %)	0.9038 (13.6 %)
`sum3`	1332.5	821.6	310.6	0.00465 (6.8 %)	0.00478 (8.7 %)	0.9717 (11.0 %)
	2614.3	2103.3	1592.6	0.16648 (5.1 %)	0.15386 (7.6 %)	1.0820 (9.0 %)

Table 4.2: Observed escape peak ratios in the test spectra with relative uncertainties. E_p, E_s and E_d are the energies of the photopeak, single escape peak and double escape peak, respectively. The single escape peak corresponding to $E_p = 1299\,\mathrm{keV}$ was not found in the first two spectra.

the exception of the last spectrum whose dead time had not been recorded. Source-detector distances were measured from the end-cap to the pointlike sources on the symmetry axis. Background nuclides were omitted from the nuclide contents, although ^{40}K and thorium and uranium daughters can be identified from the spectra. ^{208}Tl is an exception as it has a prominent gamma-ray at 2614 keV that was used in the escape peak studies. Instead of its activity, Table 4.1 presents the average counting rate of the 2614 keV gamma-rays in these measurements, because activity is not a relevant quantity for a background nuclide.

The spectra were analyzed with Sampo 90 version 3.50 [14]. Since single escape peaks are wider than photopeaks, their areas cannot usually be determined accurately using the shape calibration obtained from photopeaks. For the most accurate results, the peak shape should be calibrated separately for escape peaks. In these analyses, the most consistent results were obtained when both single and double escape peaks were fitted with a free shape, i.e., the shape calibration was only used for photopeaks. The single-escape-to-photopeak ratios R_1, double-escape-to-photopeak ratios R_2 and single-to-double-escape-peak ratios R' calculated on the basis of these analyses are presented in Table 4.2.

In Table 4.2, the value of R' seems not to be very constant as a function of energy: the weighted averages of $R'(1333\,\mathrm{keV})$, $R'(1408\,\mathrm{keV})$ and $R'(2614\,\mathrm{keV})$ are 0.79, 0.63 and 0.98, respectively. The weighted average of all R'-values is $\bar{R}' = 0.88 \pm 0.02$. It is comparable within uncertainties to the value 0.84 ± 0.05 calculated with the empirical equation of Ref. [100], so the observed discrepancy is most probably due to the small significance of our escape peaks at the two lower energies. Similarly, the values R_1 and R_2 at energies 1333, 1408 and 2614 keV do not support the hypothesis of being detector-dependent constants too well, but the hypothesis cannot either be proven false due to their relatively large uncertainties.

In order to test the escape peak prediction method of Ref. [98], i.e., $R_2\epsilon_p = s\kappa$, the values of κ and $R_2\epsilon_p$ are presented in Table 4.3. The slopes of the straight lines that go through

Spectrum	E_p (keV)	κ	$R_2\epsilon_p(E_p)$	Slope s
`esc1`	1299.1	0.01037	1.1597×10^{-5}	0.0011183
	1332.5	0.01362	1.4495×10^{-6}	0.0001064
	1408.1	0.02258	3.3598×10^{-6}	0.0001488
	2614.3	0.31006	3.1068×10^{-5}	0.0001002
				$\bar{s}=9.906\times10^{-5}$
`esc2`	1299.0	0.01036	1.2726×10^{-5}	0.0012288
	1332.3	0.01361	1.5163×10^{-6}	0.0001114
	1407.9	0.02256	2.9829×10^{-6}	0.0001322
	2614.1	0.30999	3.0635×10^{-5}	0.00009883
				$\bar{s}=9.792\times10^{-5}$
`sum1`	1332.2	0.01360	2.1132×10^{-6}	0.0001554
	2614.0	0.30996	3.3266×10^{-5}	0.0001073
				$\bar{s}=1.063\times10^{-4}$
`sum2`	1333.0	0.01367	8.7541×10^{-6}	0.0006402
	2615.2	0.31029	1.4453×10^{-4}	0.0004658
				$\bar{s}=4.621\times10^{-4}$
`sum3`	1332.5	0.01362	8.5215×10^{-6}	0.0006255
	2614.3	0.31004	1.4915×10^{-4}	0.0004810
				$\bar{s}=4.780\times10^{-4}$

Table 4.3: Calculated slopes of $(\kappa, R_2\epsilon_p)$ plots. The slopes for single $(\kappa, R_2\epsilon_p)$ pairs were calculated assuming that the plot goes through the origin. The average slope $\bar{s}$ for each spectrum is the result of a linear least squares fit of the origin and all $(\kappa, R_2\epsilon_p)$ pairs, excluding the first one in the spectra `esc1` and `esc2` since it has a photopeak component included.

the origin and each data point are also presented, as well as the slope of the line calculated as a linear least squares fit of all data points and the origin. The slope values in the first three spectra should be equal since they have been measured with the same setup, and they are indeed within 6 % of the average $\bar{s}=1.0\times10^{-4}$. Values calculated from individual peaks differ by 50 % from the average, which is quite a large difference in comparison to relative peak area uncertainties. A larger data set would be needed to investigate this. The last two spectra show individual slope values due to different measuring setups.

It can be seen from Table 4.3 that the slope calculated for $E_p = 1299\,\text{keV}$ in the spectra `esc1` and `esc2` differs substantially from the other values, but this is not an indication of any deficiency in the escape peak area prediction method. On the contrary, this result shows the power of the method: the different slope value indicates the fact that the 277 keV peak in these spectra cannot be explained as a pure double escape peak of the 1299.2 keV gamma, but it has to include some other component. In this case, only 8–9 % of the 277 keV peak area can be explained as a double escape peak and a photopeak of ^{208}Tl at 277.3 keV explains the remaining 90 % share. This interpretation is consistent with the other ^{208}Tl peaks in the spectra. The slope values calculated for $E_p = 1299\,\text{keV}$ were excluded from the best estimates for the slope for this reason.

E_p (keV)	E_s or E_d	`esc1`	`esc2`	`sum1`	`sum2`	`sum3`
$\bar{s}$		1.0×10^{-4}	1.0×10^{-4}	1.0×10^{-4}	4.6×10^{-4}	4.8×10^{-4}
$\bar{R}'$		0.88	0.88	0.88	0.88	0.88
1299.2	277.2	8.9 %	8.1 %			
1332.5	310.5	94 %	90 %	64 %	72 %	77 %
1332.5	821.5	123 %	89 %	67 %	102 %	69 %
1408.0	386.0	67 %	76 %			
1408.0	897.0	128 %	95 %			
2614.4	1592.4	100 %	101 %	93 %	99 %	100 %
2614.4	2103.4	83 %	87 %	92 %	96 %	81 %

Table 4.4: Explanation levels of the escape peaks in the test spectra, when an explanation level of 100 % is assumed for the corresponding photopeaks. The 277.2 keV peak is clearly seen to contain a photopeak component in addition to the double escape component.

4.5.3 Conclusion

Table 4.4 shows the explanation levels of the escape peaks when the explanation levels of the corresponding photopeaks are assumed to be 100 %. The explanations vary typically between 70 %–130 %, which is acceptable for nuclide identification purposes. It can be concluded that escape peak area prediction works well with the method of Ref. [98]. We can obtain a rough estimate for all escape peak areas if only one photopeak with the corresponding escape peaks is available. Of course, the accuracy of the prediction increases with an increasing number of peak triplets. However, the different shape of single escape peaks should be taken into account, or otherwise, there may be some loss in accuracy.

The two parameters needed in the method, s and R', can be determined either externally, i.e., with the help of calibration spectra or other representative spectra with escape peaks, or internally for each spectrum to be identified. A pitfall in the internal determination is that escape peak candidates may contain significant photopeak or sum peak components, thus distorting the peak area ratios. Therefore, it can only be recommended in cases where the spectrum contents are sufficiently well known in advance, either from a provisional identification or from identification results of similar spectra in a measurement series.

The method has been in production use in an air filter analysis pipeline for CTBTO radiation monitoring network (see Sec. 2.2.4). These samples always contain ^{208}Tl with its 2614 keV peak and the corresponding escape peaks. By inspection of our analysis results of over 600 spectra measured with 23 different detectors in May 2004, the conclusion above can be confirmed: the scatter in the explanation level of 1592 and 2103 keV escape peaks is only slightly larger than that of the 2614 keV photopeak. A useful indicator is the standard deviation of the explanation level divided by the corresponding average $T = \sigma_x/\bar{x}$ calculated for each detector separately. The median values of T for 23 detectors is 10 % for the photopeak and 12–13 % for the escape peaks. This difference is acceptable, especially taking the smaller size of escape peaks into account.

4.6 Random Sum Peaks

4.6.1 Calculation Method

A random sum pulse results from a detection of two gamma-rays within the resolving time of the spectrometer. The actual summing of pulses occurs in the ADC as a result of pileup of input pulses from the linear amplifier.[c] The amplitude of the summed pulse determined by the ADC depends on the time difference between the individual input pulses. If the pulses come within a very short interval, clearly shorter than the risetime of the pulses, the amplitude of the sum pulse corresponds to the sum of the individual pulses and a pulse will be recorded in the random sum peak. Otherwise, the sum pulse will add to the more or less smooth continuum below the sum peak.

The counting rate of random sum pulses n_r is proportional to the product of counting rates of the summing gamma-rays n_1 and n_2:

$$n_r = 2\tau n_1 n_2 \ . \tag{4.16}$$

Counting rates n_i ($i \in \{1, 2, r\}$) are defined as the photopeak area A_i divided by the live measuring time t_l. The proportionality constant τ has the dimension of time so that it can be interpreted as the maximum time interval between the summing pulses in order to create a completely summed pulse that is recorded in the random sum peak. The factor 2 emphasizes the fact that pulse 2 may arrive before or after pulse 1. This is analogous to the equation for chance coincidence pulse rate in a simple coincidence measurement [22, p. 636].

Using this model, the prediction of random sum peak areas requires only that the proportionality constant τ is known. However, it is relatively difficult to determine, because its dependence on the energies of the summing gamma-rays, their counting rates as well as the characteristics of detector, preamplifier, amplifier and MCA may be complicated. To a first approximation, we will assume that τ does not depend on gamma-ray energies or counting rates but only on the length of the amplifier pulses. Thus, the value of τ should be a constant in a measured spectrum and it should be linearly proportional to the amplifier shaping time constant when the spectrometer remains otherwise unaltered.

We must restrict the study to the cases where the total counting rate is not too high, because photopeak shapes tend to become distorted when the dead time percentage increases. Random sum peaks also have a decent shape when dead time is 20 % or less, but at higher counting rates and dead times, their low energy side becomes excessively strong and the determination of the peak area very inaccurate. This behavior is naturally dependent on the spectrometer in use, e.g., if a pileup rejector is used or not, and on the spectrum analysis method.

---

[c] Applicability of this method for modern spectrometers using digital signal processing (DSP) has not been evaluated.

4.6.2 Experimental Results

All random sum peak candidates in our five test spectra for random summing (see Table 4.1) are presented in Table 4.5 [V]. The spectra were analyzed with SAMPO 90 version 4.05 [14].

First of all, these data show that the energies of pure random sum peaks determined by SAMPO 90 tend to be slightly lower than the summed energy of the corresponding photopeaks. In our measurements this difference varies in the range $-1.4 \ldots -0.4$ keV. The difference is probably due to the asymmetric shape of random sum peaks that slightly distorts the centroid computation.

Table 4.5 presents the proportionality factor τ calculated from Eq. 4.16 for each random sum peak. The calculated values are of the order of 1 μs, typically 15 % of the shaping time constant used in the measurement (see Table 4.1). Our assumption was that this factor should be a constant in each spectrum, but the results do not seem to support this assumption. However, some of the worst deviations can be explained by the fact that all peaks in the table are not pure random sum peaks. The peaks with a label C in the last column contain a significant share of true coincidence summing and the peaks with a label P contain photopeak components.

The 2505 keV peaks in the spectra `sum1`, `sum2` and `sum3` are mainly produced by true coincidence summing of ^{60}Co gamma-rays. Using Eqs. 4.20 and 4.21, we can obtain an estimate for the area of the coincidence sum peak at this energy: approximately 80 %, 79 % and 85 % of the 2505 keV peak area is explained by coincidence summing in the spectra `sum1`, `sum2` and `sum3`, respectively. Thus, the random sum peak explanation levels of 5.5 %, 5.5 % and 2.1 % are acceptable as they make the total explanation of the peaks approach 100 %.

Similarly, the 1235 keV peak in the spectrum `meas14` is mainly produced by true coincidence summing of ^{116m}In gamma-rays. Coincidence summing explains a 46 % share of the peak and the two random summing channels 28 % in total.

A photopeak component of ^{133}Xe explains 63 % of the 161 keV peak and a photopeak component of ^{135}Xe explains 117 % of the 1063 keV peak in the spectrum `stuk6`. These peak shares were taken from identification results where nuclide activities were calculated in a least squares sense. When the calculated random sum explanation levels are added to these explanations, the total peak explanation percentages go slightly farther off the ideal 100 %, but the change is not very substantial. On the other hand, the 1612 keV peak with label U remains without a sufficient explanation, but the peak is so small that it may in fact be spurious.

The assumption of a linear dependence between τ and the amplifier pulse length seems to be supported by the experimental data. The average values $\bar{\tau}$ in the spectra `sum1` and `sum2` (shaping time constant 6 μs) are approximately equal and approximately three times larger than that in the spectrum `sum3` (shaping time constant 2 μs). Although there is only one pure random sum peak in the spectrum `meas14`, the calculated τ is in accordance with the $\bar{\tau}$-values of spectra `sum1` and `sum2` within a factor of two. This is an expected result, since these three spectra were measured with the same spectrometer using the same shaping time constant. On the other hand, the $\bar{\tau}$-value of the spectrum `stuk6`, calculated as an average of five τ-values, is a factor of two smaller than the others measured with the same shaping

Spectrum	E_r (keV)	E_1 (keV)	E_2 (keV)	$E_r - E_1 - E_2$ (keV)	τ (μs)	(σ_τ/τ)	Random sum peak explanation percentage	Total peak explanation percentage
`sum1`	720.2	59.4	661.5	-0.736	1.354	(11.5 %)	58	58
	1322.4	661.5	661.5	-0.663	0.195	(11.7 %)	405	405
	1391.2	59.4	1332.2	-0.505	1.510	(22.2 %)	52	52
	1833.7	661.5	1173.0	-0.857	0.399	(19.2 %)	198	198
	1992.5	661.5	1332.2	-1.317	0.485	(17.0 %)	163	163
	2505.0	1173.0	1332.2	-0.297	14.246	(4.9 %)	5.5	86 C
					$\bar{\tau} = 0.79$			
`sum2`	720.8	59.6	661.9	-0.648	1.137	(3.2 %)	73	73
	1232.8	59.6	1173.7	-0.509	1.331	(10.3 %)	62	62
	1323.1	661.9	661.9	-0.648	0.264	(2.6 %)	315	315
	1391.8	59.6	1333.0	-0.724	1.182	(11.6 %)	70	70
	1834.5	661.9	1173.7	-1.046	0.534	(4.9 %)	155	155
	1993.9	661.9	1333.0	-0.955	0.530	(4.8 %)	157	157
	2506.2	1173.7	1333.0	-0.485	15.081	(1.9 %)	5.5	85 C
					$\bar{\tau} = 0.83$			
`sum3`	720.6	59.5	661.6	-0.544	0.364	(7.2 %)	80	80
	1232.1	59.5	1173.2	-0.646	0.405	(17.4 %)	72	72
	1322.7	661.6	661.6	-0.574	0.0755	(5.7 %)	384	384
	1391.1	59.5	1332.5	-0.954	0.534	(15.6 %)	54	54
	1833.5	661.6	1173.2	-1.386	0.186	(8.5 %)	156	156
	1993.4	661.6	1332.5	-0.686	0.154	(10.2 %)	188	188
	2505.2	1173.2	1332.5	-0.534	14.119	(1.8 %)	2.1	87 C
					$\bar{\tau} = 0.29$			
`meas14`	1234.6	138.0	1096.8	-0.192	10.255	(18.2 %)	14	74 C
	1234.6	416.5	818.2	-0.109	9.788	(18.2 %)	14	74 C
	1708.9	416.5	1293.0	-0.595	1.370	(16.5 %)	102	102
					$\bar{\tau} = 1.4$			
`stuk6`	160.9	80.8	80.8	-0.765	0.549	(5.8 %)	78	141 P
	280.1	30.5	250.0	-0.425	0.670	(7.4 %)	64	64
	330.3	80.8	250.0	-0.478	0.477	(7.1 %)	90	90
	499.0	250.0	250.0	-0.963	0.191	(6.1 %)	225	225
	688.4	80.8	608.3	-0.691	0.487	(9.5 %)	88	88
	857.1	250.0	608.3	-1.172	0.314	(7.2 %)	137	137
	1062.6	250.0	812.8	-0.110	11.769	(8.4 %)	3.7	121 P
	1612.2	80.8	1530.1	$+1.258$	13.483	(8.4 %)	3.2	3.2 U
					$\bar{\tau} = 0.43$			

Table 4.5: Random sum peak data in the test spectra. E_r is the sum peak energy, E_1 and E_2 are the energies of the summing gamma-rays, and τ is the proportionality factor (with uncertainty) calculated from Eq. 4.16. The unweighted averages $\bar{\tau}$ have been calculated excluding the peaks with significant true coincidence summing (label C), photopeak (label P) or unknown (label U) components. The last two columns present the peak explanation percentage by the random sum peak model (actually $\bar{\tau}/\tau$) and the total explanation percentage of the peak as given by the Shaman analysis, respectively.

time constant. This difference is likely to be significant, but this is also expected as the spectrum `stuk6` was measured with another spectrometer. It is reasonable that τ does not solely depend on the shaping time constant but also on other settings. For example, the source-detector distance is different from the other measurements and it may have some effect, although the default assumption is that only count rate affects random summing, not the solid angle subtended by the detector. However, the source-detector distance may have an effect on the time behavior of pulse generation in the detector and thus on summing of pulses. Different dependences should be investigated in a more extensive study.

Even when the explanatory remarks above are taken into account, the calculated τ-values vary significantly from one peak to another in each spectrum. When only the pure random sum peaks are taken into account, the ratio of the largest value of τ to the smallest one is 6.9, 5.0, 7.1, and 3.5 in the spectra `sum1`, `sum2`, `sum3`, and `stuk6`, respectively. Thus, the assumption of the constant proportionality factor τ does not seem to be strictly correct, but if a qualitative identification of random sum peaks is sufficient, a constant value for τ can be used in spectrum interpretation.

Some of the variation of τ-values can probably be explained with the slightly distorted shape of random sum peaks, rendering their area determination more inaccurate than usual, but this cannot be the only reason. On the other hand, it is evident from Table 4.5 that the values of τ have no clear correlation with energy. No clear correlation of τ-values with the product of counting rates $n_1 n_2$ was either found.

However, a clear positive correlation between τ and $|E_1 - E_2|$ could be established. The correlation coefficient of these two variables was 0.79, 0.84, 0.89 and 0.36 in spectra `sum1`, `sum2`, `sum3` and `stuk6`, respectively, omitting the peaks with true coincidence and photopeak components. Fitting a straight line to each data set leads to the following dependences:

`sum1`: $\tau = 0.154 + 0.00104 \times |E_1 - E_2|$

`sum2`: $\tau = 0.266 + 0.000811 \times |E_1 - E_2|$

`sum3`: $\tau = 0.0489 + 0.000342 \times |E_1 - E_2|$

`stuk6`: $\tau = 0.344 + 0.000328 \times |E_1 - E_2|$

However, the uncertainties of the constants are quite large due to a substantial scatter in the data points.

This result indicates that the MCA can separate individual pulses from the signal chain the better the smaller the size difference is. An explanation for this phenomenon would probably require deeper knowledge on electronics.

Coming back to the initial assumption that τ depends linearly on the shaping time constant of the spectrometer, we pool the first two spectra to represent a shaping time $6\,\mu$s and the third spectrum to represent a shaping time $2\,\mu$s. The spectrum `stuk6` has been measured with a different spectrometer, so it is not comparable to the others. Observing the limited accuracy obtainable with these data on one hand and the relative insensitivity of final results to actual parameter values on the other hand, we select the best estimates for the spectra with shaping time $6\,\mu$s to be $\tau_0 = 0.15\ \mu$s and $\eta = 0.001\ \mu$s/keV and 1/3 of these parameters for spectra with shaping time $2\,\mu$s.

Spectrum	E_r (keV)	E_1 (keV)	E_2 (keV)	τ_0 (μs)	η (μs/keV)	Random sum peak explanation percentage	Total peak explanation percentage
`sum1`	720.2	59.4	661.5	0.15	0.001	56	56
	1322.4	661.5	661.5	0.15	0.001	77	77
	1391.2	59.4	1332.2	0.15	0.001	94	94
	1833.7	661.5	1173.0	0.15	0.001	166	166
	1992.5	661.5	1332.2	0.15	0.001	169	169
	2505.0	1173.0	1332.2	0.15	0.001	2.2	83 C
`sum2`	720.8	59.6	661.9	0.15	0.001	66	66
	1232.8	59.6	1173.7	0.15	0.001	95	95
	1323.1	661.9	661.9	0.15	0.001	57	57
	1391.8	59.6	1333.0	0.15	0.001	120	120
	1834.5	661.9	1173.7	0.15	0.001	124	124
	1993.9	661.9	1333.0	0.15	0.001	155	155
	2506.2	1173.7	1333.0	0.15	0.001	2.1	83 C
`sum3`	720.6	59.5	661.6	0.05	0.0003	63	63
	1232.1	59.5	1173.2	0.05	0.0003	95	95
	1322.7	661.6	661.6	0.05	0.0003	66	66
	1391.1	59.5	1332.5	0.05	0.0003	81	81
	1833.5	661.6	1173.2	0.05	0.0003	109	109
	1993.4	661.6	1332.5	0.05	0.0003	163	163
	2505.2	1173.2	1332.5	0.05	0.0003	0.7	86 C
`meas14`	1234.6	138.0	1096.8	0.15	0.001	11	62 C
	1234.6	416.5	818.2	0.15	0.001	5.6	62 C
	1708.9	416.5	1293.0	0.15	0.001	75	75
`stuk6`	160.9	80.8	80.8	0.30	0.0003	55	117 P
	280.1	30.5	250.0	0.30	0.0003	55	55
	330.3	80.8	250.0	0.30	0.0003	74	74
	499.0	250.0	250.0	0.30	0.0003	157	157
	688.4	80.8	608.3	0.30	0.0003	94	94
	857.1	250.0	608.3	0.30	0.0003	130	130
	1062.6	250.0	812.8	0.30	0.0003	4.0	121 P
	1612.2	80.8	1530.1	0.30	0.0003	5.5	5.5 U

Table 4.6: Random sum peak explanation levels with the two-parameter model. The data are otherwise the same as in Table 4.5, but the columns τ_0 and η show the values applied for each spectrum. The scatter in the peak explanation percentages is seen to decrease significantly from that obtained with the constant-τ model.

Table 4.6 presents the explanation percentages of random sum peaks obtained with a two-parameter model $\tau = \tau_0 + \eta|E_1 - E_2|$ using the selected parameter values. A clear improvement is obtained in comparison to the constant-τ model used in Table 4.5. This is indicated by the fact that the maximum-to-minimum explanation ratio for the pure random sum peaks decreased to 3.0, 2.7, 2.6 and 2.9 in the spectra `sum1`, `sum2`, `sum3`, and `stuk6`, respectively. The average of pure random sum peak explanation percentages is 102 % and its standard deviation 39 %, when the corresponding parameters for the constant-τ model were 146 % and 103 %.

4.6.3 Conclusion

We suggested a conceptually simple method for the prediction of random sum peak areas and evaluated it with a few test spectra. There is only one parameter τ in the model and in this work it was tested whether this parameter was a constant in a measured spectrum or not. When a constant parameter was used, the method was found to predict sum peak areas correctly within a factor of three. This accuracy is sufficient only for qualitative interpretation of a spectrum.

However, when the parameter τ was found to be correlated with the energy difference of the summing gamma-rays $|E_1 - E_2|$, a linear two-parameter model for τ was suggested and evaluated. It was found to improve the accuracy of the predictions to within a factor of two or slightly better. This accuracy is still worse than for other spectrum peaks, but it may be acceptable for some applications. However, applicability of the method was practically evaluated only for one spectrometer. As this work shows, introduction of the method requires a thorough investigation of the random summing behavior in each spectrometer separately.

4.7 True Coincidence Sum Peaks

4.7.1 Calculation Method

Estimation of Coincidence Sum Peak Areas

True coincidence summing occurs when two or more gamma-rays emitted in a cascade by a decaying nuclide are detected within the resolving time of the spectrometer. The extent of coincidence summing depends on the solid angle subtended by the detector, i.e., the source-detector distance and the area of the detector front face. In addition to producing coincidence sum peaks, coincidence summing modifies the ratios of photopeak areas when compared to a reference database. Coincidence corrections for photopeak areas can be calculated if the decay schemes of the measured radionuclides, the photopeak efficiency and the total efficiency are known, as has been demonstrated by Andreev et al. in Ref. [101]. Their method has been further developed and applied in practice by other authors, e.g., in Refs. [102] – [104]. Other calculation methods for true coincidence correction have also been presented, e.g., in Refs. [105] – [110].

Calculation of coincidence sum peak areas has not been discussed in the previous publications, with the exception of Refs. [108] – [110] where a matrix notation has been applied. However, we have extended the method of Andreev et al. to also calculate sum peak areas [II, IV]. Our principle is simple: the crossover transitions missing from the decay scheme of the nuclide due to their low or zero emission probability are added to the scheme with the sum energy and a small initial emission probability. An initial value of zero cannot be used, because it would lead to divisions by zero. If the summing-in correction of a crossover transition is significant, its apparent emission probability will grow in the coincidence correction calculations, which in turn decreases the apparent emission probabilities of the cascade gamma-rays. The apparent emission probability of the crossover gamma-ray is obtained

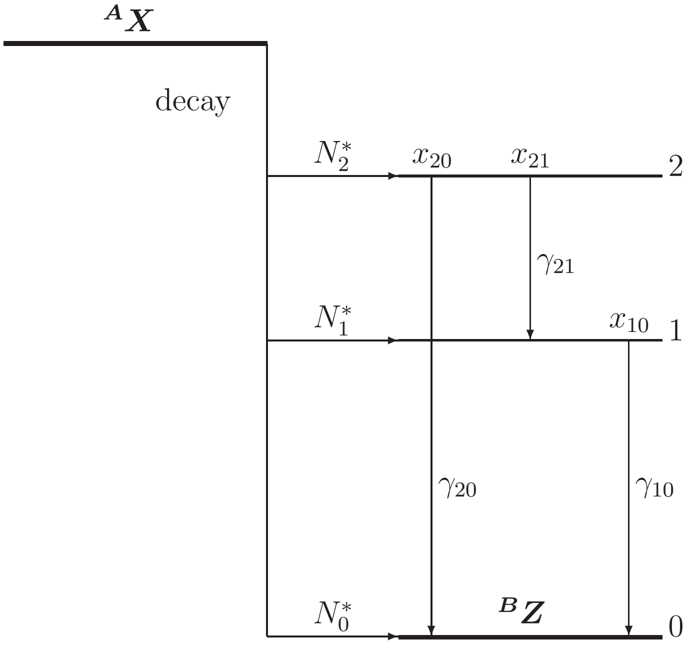

Figure 4.10: A simple decay scheme for illustrating the calculation of coincidence correction coefficients. As a result of the decay of ^{A}X, three different gamma-rays are emitted with transition probabilities x_{ij}. The N_i^* values are the decay branching ratios to different energy levels.

by subtracting the initial emission probability from the calculation result. The crossover transitions added to the decay scheme affect correction factors of the cascade gamma-rays, but this distortion can be made insignificant by choosing a sufficiently low initial emission probability (e.g., 0.001 %).

The method can be illustrated with the help of the simple decay scheme of Fig. 4.10. We neglect the effect of angular correlation and assume that the half-life of the energy level 1 is small in comparison to the resolving time of the spectrometer. In this example, we will also neglect internal conversion[d] although it is included in our complete calculation model. If we set the branching ratios of the two lowest states N_0^* and N_1^* as well as the crossover transition probability x_{20} to zero, the values of N_2^*, x_{21} and x_{10} have to be 1. This corresponds to the decay scheme of ^{60}Co, for example, to a sufficient accuracy. Applying the correction formulas of Ref. [101], we can derive the observed peak areas S_{ij} to be:

$$\begin{aligned}
S_{21} &= At_l N_2^* x_{21} \epsilon_{p21} x_{10}(1-\epsilon_{t10}) = At_l \epsilon_{p21}(1-\epsilon_{t10}) = S_{21}^*(1-\epsilon_{t10})\ , && (4.17)\\
S_{10} &= At_l (N_1^* + N_2^* x_{21}(1-\epsilon_{t21})) x_{10} \epsilon_{p10} = At_l \epsilon_{p10}(1-\epsilon_{t21}) = S_{10}^*(1-\epsilon_{t21})\ , && (4.18)\\
S_{20} &= At_l N_2^* x_{20} \epsilon_{p20} + At_l N_2^* x_{21} \epsilon_{p21} x_{10} \epsilon_{p10} = At_l \epsilon_{p21} \epsilon_{p10}\ , && (4.19)
\end{aligned}$$

where

A = average activity of the nuclide during the measurement,

---

[d] Internal conversion, i.e., emission of electrons, is a competing reaction for gamma-ray emission. It is included in the calculation by multiplying all efficiencies with $1/(1+\alpha_{ij})$ where α_{ij} is the conversion coefficient that is specific for transition $i \rightarrow j$.

t_l = live time of the measurement,
S^*_{ij} = peak area without coincidence summing for transition $i{\rightarrow}j$,
N^*_i = branching ratio to level i,
x_{ij} = probability of transition $i{\rightarrow}j$ (incl. both gamma-ray emission and internal conversion),
ϵ_{pij} = photopeak efficiency at the energy of transition $i{\rightarrow}j$,
ϵ_{tij} = total efficiency at the energy of transition $i{\rightarrow}j$.

It can be seen that the correction factors of the cascade peak areas S_{21} and S_{10} are of the typical form $(1 - \epsilon_t)$. The coincidence sum peak area S_{20}, on the other hand, is proportional to the product of the photopeak efficiencies at the cascade energies. The ratios of the cascade peak areas to the sum peak area are:

$$\frac{S_{21}}{S_{20}} = \frac{1-\epsilon_{t10}}{\epsilon_{p10}} \approx \frac{1}{\epsilon_{p10}}\,, \tag{4.20}$$

$$\frac{S_{10}}{S_{20}} = \frac{1-\epsilon_{t21}}{\epsilon_{p21}} \approx \frac{1}{\epsilon_{p21}}\,, \tag{4.21}$$

where the approximations are valid when $\epsilon_t \ll 1$. In more complicated decay schemes the equations also become more complicated, but the peak areas can still be calculated without approximations using the method of Ref. [101]. Nevertheless, there are two omissions in this coincidence sum peak estimation method:

1. It does not estimate sum peak areas corresponding to coincidence summing of non-consecutive transitions, e.g., transitions 3→2 and 1→0 when transition 2→1 exists. Typically these sum peaks are weak, but they are occasionally observed.
2. It does not estimate sum peak areas corresponding to coincidence summing of a gamma-ray and an X-ray or of two X-rays. These are observed in cases where strongly converted transitions exist, leading to X-rays in abundance.

The latter deficiency with sum peaks including X-ray components is related to the fact that coincidence corrections of X-ray lines cannot be calculated by the method of Ref. [101]. Although summing of gamma-rays with X-rays is accounted for from the summing-out perspective, both for X-rays following internal conversion and those following electron capture, corrections in X-ray emission probabilities are not covered by the method. X-ray summing phenomena have been investigated with Monte Carlo simulation in Ref. [111] for the ^{212}Pb chain that is important in air filter samples. A general analytical formulation has also been derived in Ref. [111], but it is not reproduced here and it has not been implemented by us.

The former deficiency, summing of gamma-rays of non-consecutive transitions, can be remedied in the calculation module relatively simply by considering the issue from a detection probability perspective. Assume a decay scheme like in Fig. 4.10 but with a third excited level with a non-zero branching ratio N^*_3 and a non-zero transition probability x_{32}. Gamma-rays from the transition 3→2 would experience summing with those from transitions 2→0 and 2→1, already covered by the basic calculation method, but also with gamma-rays from transition 1→0. The latter summing corresponds to the following kind of decay histories:

1. Gamma-ray 3→2 is fully absorbed in the detector. This happens with the probability $P_3 = \epsilon_{p32}$.

2. Not even a part of gamma-ray 2⟶1 is absorbed in the detector. This happens with the probability $P_2 = 1 - \epsilon_{t21}$.
3. Gamma-ray 1⟶0 is fully absorbed in the detector. This happens with the probability $P_1 = \epsilon_{p10}$.

When the frequency of these histories is $AN_3^* x_{32}$, we obtain the following sum peak area with a measurement live time t_l:

$$S_{3210,approx} = At_l N_3^* x_{32} P_3 P_2 P_1 = At_l N_3^* x_{32} \epsilon_{p32} (1 - \epsilon_{t21}) \epsilon_{p10} \ . \tag{4.22}$$

However, this is only the first approximation that omits the fact that all transitions from level 2 do not lead to level 1. Additionally, internal conversions of all three transitions modify the probabilities presented above. Taking these adjustments into account, we obtain the following equation for the sum peak area:

$$S_{3210} = At_l \times N_3^* \frac{x_{32}\epsilon_{p32}}{1+\alpha_{32}} \times x_{21} \left(1 - \frac{\epsilon_{t21}}{1+\alpha_{21}}\right) \times \frac{\epsilon_{p10}}{1+\alpha_{10}} \ , \tag{4.23}$$

where α_{ij} are the total conversion coefficients of each transition. This extension for the sum peak calculation has not yet been implemented in SHAMAN.

Manipulation of Decay Schemes due to Metastable States

Decay schemes are needed in coincidence correction calculations. Traditionally, schemes have been compiled manually from data collected from suitable sources, e.g., Refs. [112] – [115]. It has been customary to omit the weakest transitions from the schemes because the corresponding gamma-rays are practically never seen in measured spectra. Paradoxically, the weakest gamma-ray lines would usually experience the strongest corrections. Automated scheme compilation has been attempted, e.g., in Ref. [116], but manual inspection is still required due to inconsistencies in basic nuclear databases.

Another reason for compiling decay schemes manually is that they typically require some manipulations to produce correct results. Ref. [102] presents a method to account for coincidences with annihilation gamma-rays following β^+-decay. This is accomplished by adding a virtual level to the decay scheme 511 keV above the true level. Coincidences with X-rays following EC-decay can also be compensated for by similar decay scheme manipulation [103, 104], but this can be handled inside the calculation algorithm.

Metastable states in nuclear decay schemes also raise a need for scheme manipulation, because a long-lived state prevents a simultaneous detection of a gamma-ray leading to this state and any other gamma-ray initiating from this state. This behavior can be accomplished by replacing the gamma-ray transitions leading to the metastable state by gamma-ray transitions from suitable virtual levels to the ground state. An example involving the β^--decay scheme of ^{91}Sr is presented in Fig. 4.11a where weak transitions have been omitted [V]. In this case, the levels 3 and 4 have to be manipulated so that coincidences $\gamma_{41}+\gamma_{10}$ and $\gamma_{31}+\gamma_{10}$ are prohibited but coincidences $\gamma_{32}+\gamma_{20}$ are allowed. This can be done by virtually moving the level 4 downwards and splitting the level 3 into two levels as shown in Fig. 4.11b.

This kind of manipulation can usually be done relatively easily, but it would be better suited for the calculation algorithm. This would ensure that modifications are made consistently

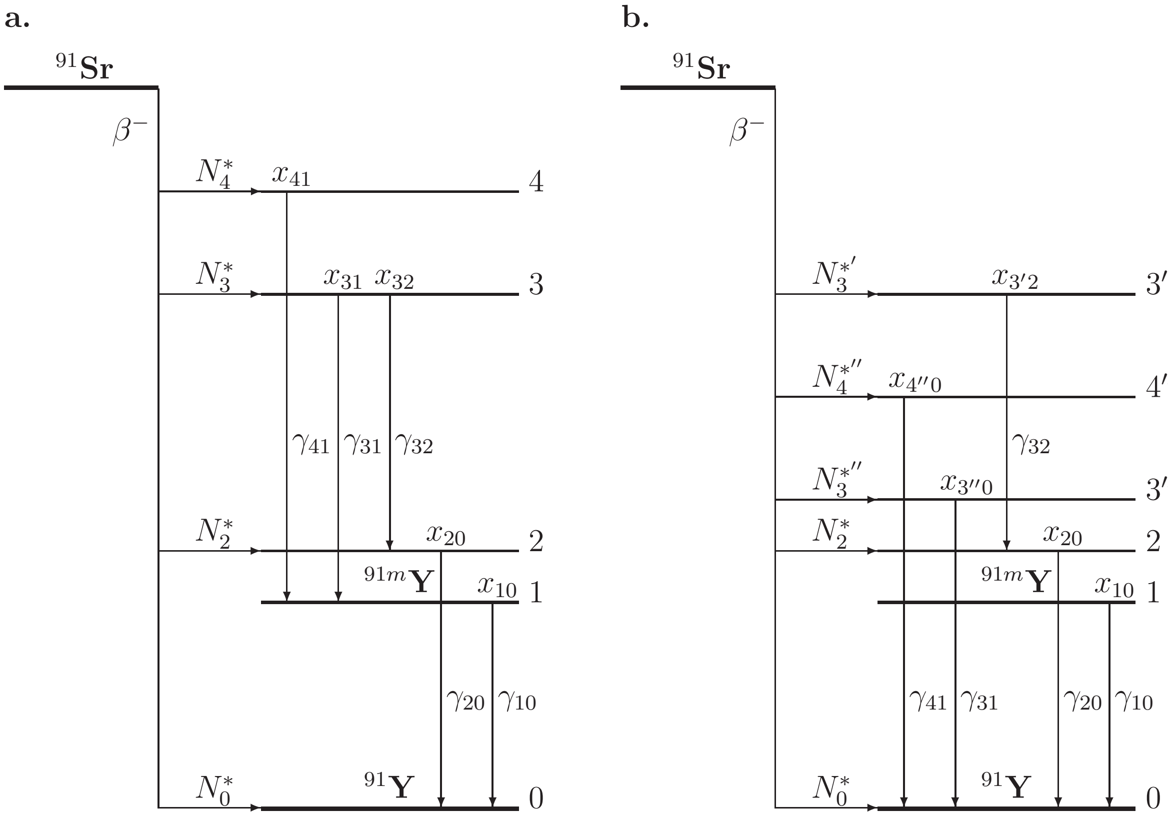

Figure 4.11: An example of decay scheme manipulation to compensate for a metastable state: **a.** the original decay scheme [114], **b.** the modified decay scheme. The new values for branching ratios and transition probabilities must be chosen as follows: $N_4^{*''} = N_4^*$, $N_3^{*''} = N_3^* x_{31}$, $N_3^{*'} = N_3^* x_{32}$, and $x_{4''0} = x_{3''0} = x_{3'2} = 1$.

for all radionuclides and for all transitions. The modifications related with annihilation gamma-rays following β^+-decay should also be made in the calculation algorithm instead of manipulating the decay scheme data file. The current practice in Shaman is to make the modifications manually.

One question worth consideration is: which is the longest half-life of an energy level that does not prevent coincidence summing? This question is important even if we are only interested in coincidence corrections of cascade gamma-rays, but for a reliable prediction of coincidence sum peak areas its effect is crucial: an omission of a metastable state will lead to crudely overestimated sum peak areas.

An estimate for the critical half-life is obtained from the random summing studies of Sec. 4.6. It was concluded that for a full summing, the maximum interval between the summing pulses is of the order of 0.1–1 μs depending on the spectrometer and its setup. This must also be true for summing-in correction in coincidence summing. For a summing-out correction, on the other hand, only a partial summing of amplifier pulses is required, and therefore, the longest half-life of an energy level that does not prevent summing must be substantially longer. It is likely to be of the order of the risetime of the amplifier pulses, typically 5–15 μs or about one half of the pulse length.

It can be concluded that decay scheme manipulation is required when nuclear states with half-lives of $0.1\,\mu$s or longer exist. However, the half-life limit is not a sharp one, it is different for summing-in and summing-out corrections, and it varies from one measurement to another. It can be stated that half-lives in a "transition interval" 0.1–$100\,\mu$s are likely to complicate coincidence correction calculations.

The transition interval for the summing-in effect can be predicted theoretically. Assume that coincidence summing of two gamma-rays is possible if their time interval is shorter than a value τ. If a nuclear state has a half-life $T_{1/2}$ and a decay constant $\lambda = \ln 2/T_{1/2}$, the share of decays before the time τ has elapsed is

$$p = \frac{\text{decays during } [0,\tau]}{\text{all decays}} = 1 - e^{-\lambda\tau} = 1 - 2^{-\tau/T_{1/2}} \ . \tag{4.24}$$

When $T_{1/2} = \tau$, 50 % of the decays occur so that full summing is possible. This would mean that the predicted sum peak area is twice the observed one if the metastable state was unaccounted for. A transition from $p = 90\,\%$ to $p = 10\,\%$ occurs when $T_{1/2}$ increases from 0.3τ to 7τ. In other words, full summing would not be significantly affected by half-lives below 0.3τ and it would be almost completely prohibited by half-lives above 7τ.

A similar transition should occur in partial summing, only the value of τ would be larger. This transition is generally more important than the transition in full summing, because summing-out corrections are typically larger (of the form $1/(1-\epsilon_t)$) than summing-in corrections (of the form $1/(1+\epsilon_{p1}\epsilon_{p2})$). However, pure sum peaks are most severely affected by long-lived states, as they gradually vanish when the half-life of the intermediate state increases.

Nuclides with metastable states in the transition interval should be measured in close geometries in order to verify the existence of the predicted transition and to determine its limits quantitatively. An example nuclide is ^{187}W with a $0.5553\,\mu$s state between the abundant 72.0 keV and 479.5 keV transitions [114]. According to the discussion above, this half-life should not affect summing-out corrections of the cascade gamma-rays, but the summing-in effect observed in the 551.5 keV crossover gamma-ray should be smaller than predicted without decay scheme manipulation. With the help of ^{187}W and a couple of other nuclides with intermediate state half-lives in the transition interval, e.g., ^{67}Ga and ^{169}Yb, the discussed effects could be studied and quantified.

From a practical point of view, it is fortunate that state half-lives within the transition interval are not very common. This can be seen in a complete ENSDF database [113], from which about 80 000 gamma-ray lines can be extracted. The database includes 30 000 nuclear level records, of which only 242 records or 0.8 % have a half-life between $0.1\,\mu$s and $100\,\mu$s. Actually, a vast majority of the nuclear levels do not have their half-lives specified, but they are likely to be levels with a half-life in the nanosecond-scale or below. Additionally, it must be noted that many level records are present in the database in more than one entry, but this should not affect relative shares. The presented figures give a clear indication that nuclides with state half-lives within the transition interval are rare.

Spectrum	Nuclide	E_c (keV)	E_s (keV)	Observed ratio r_o	Relative uncertainty of r_o	Calculated ratio r_c	Sum peak explanation	Coincidence correction of sum peak
`stuk1`	^{110m}Ag	657.8	1542.4	24.26	12.3 %	34.94	69 %	0.000
		884.7	1542.4	17.43	11.8 %	26.84	65 %	0.000
		657.8	1595.2	52.09	16.4 %	74.08	70 %	−1
		937.5	1595.2	18.81	16.1 %	27.01	70 %	−1
		884.7	1822.2	56.04	17.2 %	65.19	86 %	0.000
		937.5	1822.2	28.17	17.2 %	30.93	91 %	0.000
	^{134}Cs	604.7	1400.6	22.90	8.7 %	28.67	80 %	0.000
		795.9	1400.6	20.11	8.5 %	25.15	80 %	0.000
`stuk2`	^{106}Rh	511.9	1133.8	15.58	6.0 %	30.99	50 %	0.000
		621.9	1133.8	6.43	6.0 %	14.24	45 %	0.000
	^{134}Cs	569.3	1174.1	10.35	4.9 %	14.90	69 %	−1
		604.7	1174.1	74.44	4.9 %	104.05	72 %	−1
		604.7	1400.6	16.23	5.6 %	19.14	85 %	0.000
		795.9	1400.6	14.18	5.6 %	16.74	85 %	0.000
		604.7	1406.7	240.90	7.8 %	254.49	95 %	−1
		802.0	1406.7	19.67	7.9 %	20.93	94 %	−1
		604.7	1969.9	289.51	8.1 %	330.50	88 %	0.000
		1365.2	1969.9	13.57	8.2 %	14.63	93 %	0.000
`stuk3`	^{42}K	312.6	1837.3	5.11	11.0 %	9.73	53 %	0.000
		1524.6	1837.3	339.98	12.1 %	597.33	57 %	0.000
	^{99}Mo	140.5	880.0	69.13	11.9 %	5.24	1319 %	−1
		739.5	880.0	8.40	13.2 %	13.65	62 %	−1
		181.1	920.6	6.18	21.9 %	5.97	104 %	0.000
		739.5	920.6	11.56	22.6 %	11.29	102 %	0.000
	^{135}I	836.8	1968.3	12.61	12.7 %	13.79	91 %	0.000
		1131.5	1968.3	43.44	12.8 %	46.09	94 %	0.000
	^{134}Cs	604.7	1400.6	14.16	16.0 %	13.30	107 %	0.000
		795.9	1400.6	12.64	16.5 %	11.65	109 %	0.000
`stuk4`	^{46}Sc	889.3	2010.0	62.52	9.9 %	49.27	127 %	0.000
		1120.6	2010.0	61.86	10.0 %	49.10	126 %	0.000
`stuk5`	^{46}Sc	889.3	2010.0	43.74	20.3 %	49.27	89 %	0.000
		1120.6	2010.0	44.95	20.3 %	49.10	92 %	0.000
`stuk7`	^{132}I	667.7	1440.3	18.02	11.0 %	13.54	133 %	0.000
		772.6	1440.3	14.35	11.1 %	10.33	139 %	0.000
	^{134}Cs	604.7	1400.6	7.60	10.4 %	11.39	67 %	0.000
		795.9	1400.6	6.78	10.8 %	9.96	68 %	0.000
	^{140}La	487.0	2083.2	11.14	13.1 %	11.07	101 %	0.004
		1596.2	2083.2	23.31	13.5 %	24.22	96 %	0.004
	^{239}Np	106.1	383.7	18.48	9.6 %	23.26	79 %	0.000
		277.6	383.7	13.09	9.5 %	14.35	91 %	0.000
`stuk8`	^{99}Mo	181.1	920.6	2.60	6.6 %	4.70	55 %	0.000
		739.5	920.6	4.71	7.3 %	8.62	55 %	0.000
	^{132}I	667.7	1440.3	12.38	10.0 %	13.05	95 %	0.000
		772.6	1440.3	10.00	10.1 %	9.93	101 %	0.000
	^{134}Cs	604.7	1400.6	5.85	10.3 %	10.99	53 %	0.000
		795.9	1400.6	5.04	10.9 %	9.58	53 %	0.000
	^{208}Tl	510.8	1094.0	8.08	11.0 %	6.78	119 %	0.168
		583.2	1094.0	22.13	10.5 %	28.67	77 %	0.168
	^{239}Np	106.1	383.7	25.82	9.2 %	24.59	105 %	0.000
		277.6	383.7	18.28	9.1 %	14.94	122 %	0.000
`stuk9`	^{132}I	667.7	1440.3	10.36	18.0 %	13.60	76 %	0.000
		772.6	1440.3	7.44	18.1 %	10.36	72 %	0.000
	^{239}Np	106.1	383.7	24.40	17.1 %	23.10	106 %	0.000
		277.6	383.7	18.88	17.2 %	14.27	132 %	0.000

Table 4.7: Comparison of observed and calculated peak ratios of cascade gamma-rays (E_c) and sum gamma-rays (E_s) with zero or negligible initial emission probabilities. These correspond to coincidence correction 0 or slightly above 0, respectively. The coincidence correction value −1 indicates sum peaks that result from summing of non-consecutive gamma-rays in a cascade.

Spectrum	Nuclide	E_c (keV)	E_s (keV)	Observed ratio r_o	Relative uncertainty of r_o	Calculated ratio r_c	Sum peak explanation	Coincidence correction of sum peak
`stuk1`	^{110m}Ag	677.6	1562.3	3.51	15.0 %	5.66	62 %	0.900
		884.7	1562.3	27.33	14.0 %	42.26	65 %	0.900
	^{124}Sb	602.7	1325.5	51.55	8.4 %	48.36	107 %	0.846
		722.8	1325.5	5.23	8.6 %	4.89	107 %	0.846
	^{134}Cs	563.2	1168.0	3.75	7.0 %	3.54	106 %	0.928
		604.7	1168.0	46.03	6.9 %	44.80	103 %	0.928
		569.3	1365.2	3.29	8.4 %	3.62	91 %	0.869
		795.9	1365.2	19.62	8.3 %	21.83	90 %	0.869
`stuk2`	^{106}Rh	511.9	1128.1	42.45	7.3 %	44.18	96 %	0.943
		616.2	1128.1	1.16	11.0 %	1.46	80 %	0.943
		511.9	1562.3	60.26	9.0 %	77.22	78 %	0.665
		1050.4	1562.3	3.99	9.3 %	5.58	71 %	0.665
	^{134}Cs	563.2	1168.0	3.10	4.5 %	3.23	96 %	0.880
		604.7	1168.0	39.77	4.4 %	41.56	96 %	0.880
		569.3	1365.2	2.97	5.5 %	3.23	92 %	0.806
		795.9	1365.2	18.64	5.5 %	19.75	94 %	0.806
`stuk3`	^{133}I	529.9	1236.4	49.89	7.6 %	52.60	95 %	0.919
		706.6	1236.4	0.73	9.2 %	0.77	94 %	0.919
	^{135}I	546.6	1678.0	0.58	8.0 %	0.59	100 %	0.914
		1131.5	1678.0	1.85	9.3 %	1.84	100 %	0.914
		1124.0	2255.5	4.28	11.1 %	3.90	110 %	0.760
		1131.5	2255.5	25.85	10.8 %	24.13	107 %	0.760
	^{135}Xe	158.3	408.0	0.58	8.9 %	0.58	101 %	0.891
		249.8	408.0	270.10	6.9 %	223.03	121 %	0.891
`stuk4`	^{239}Np	106.1	334.3	7.47	21.4 %	11.23	67 %	0.917
		228.2	334.3	6.08	21.5 %	4.63	131 %	0.917
`stuk6`	^{88}Kr	196.3	2548.4	27.91	10.6 %	37.56	74 %	0.948
		2352.1	2548.4	0.70	18.2 %	0.98	71 %	0.948
	^{135}Xe	249.8	1062.4	14456.72	12.6 %	11859.89	122 %	0.541
		812.6	1062.4	10.85	12.1 %	8.34	130 %	0.541
`stuk7`	^{132}I	630.2	1442.6	4.20	17.6 %	4.78	88 %	0.838
		812.0	1442.6	2.35	18.7 %	2.08	113 %	0.838
	^{140}La	925.2	2521.4	1.45	12.1 %	1.53	95 %	0.903
		1596.2	2521.4	17.68	11.4 %	19.09	93 %	0.903
	^{239}Np	106.1	315.9	9.90	5.4 %	11.53	86 %	0.875
		209.8	315.9	2.18	5.4 %	1.69	129 %	0.875
		106.1	334.3	6.56	4.8 %	7.60	86 %	0.746
		228.2	334.3	10.03	4.4 %	3.50	286 %	0.746
`stuk8`	^{132}I	522.7	1295.1	4.67	8.5 %	3.95	118 %	0.718
		772.6	1295.1	26.23	7.0 %	20.17	130 %	0.718
		630.2	1297.9	3.50	10.5 %	5.24	67 %	0.558
		667.7	1297.9	23.75	9.4 %	43.47	55 %	0.558
	^{143}Ce	57.4	350.6	1.76	22.3 %	2.82	62 %	0.892
		293.3	350.6	6.59	18.3 %	11.11	59 %	0.892
		57.4	721.9	1.79	19.2 %	1.78	100 %	0.941
		664.6	721.9	1.72	19.2 %	0.94	183 %	0.941
	^{212}Bi	727.3	1512.7	16.19	15.1 %	16.49	98 %	0.770
		785.4	1512.7	2.22	17.5 %	2.42	92 %	0.770
	^{239}Np	106.1	315.9	11.71	5.0 %	11.93	98 %	0.878
		209.8	315.9	2.39	5.0 %	1.72	139 %	0.878
		106.1	334.3	7.64	5.0 %	7.89	97 %	0.751
		228.2	334.3	9.35	4.8 %	3.58	261 %	0.751
`stuk8`	^{239}Np	106.1	315.9	10.28	6.7 %	11.50	89 %	0.874
		209.8	315.9	2.37	7.3 %	1.69	140 %	0.874
		106.1	334.3	7.12	5.9 %	7.57	94 %	0.744
		228.2	334.3	9.25	5.9 %	3.49	265 %	0.744

Table 4.8: Comparison of observed and calculated peak ratios of cascade gamma-rays (E_c) and sum gamma-rays (E_s) with non-zero initial emission probabilities.

4.7.2 Experimental Results

A set of nine spectra obtained from STUK, denoted as `stuk1`…`stuk9`, was used to verify the presented method for estimating coincidence sum peak areas. The spectra had been measured with different detectors from voluminous environmental and reactor samples in a close geometry. Semi-empirical total efficiency calibrations were available for these spectra, which made the expected accuracy of the coincidence summing calculations sufficient for quantitative conclusions.

The ratio of a cascade gamma-ray to the corresponding sum gamma-ray is a sensitive indicator of eventual deficiencies in the method. If the calculation were somehow biased, peak areas of cascade gamma-rays would not decrease correctly and the sum peak areas would not increase correctly, and thus, their ratio would be distorted quadratically. The accuracy of the sum peak prediction method was verified by comparing the calculated and observed peak area ratios. This could most readily be done by calculating the ratios of coincidence corrected emission probabilities (predicted ratios) and the ratios of emission rates (observed ratios), i.e., peak areas corrected for efficiency. These values can be compared directly and they should agree within uncertainties.

The results of the comparison are shown in Table 4.7 that includes pure or almost pure sum peaks in the test spectra, and in Table 4.8 that includes sum peaks with a significant contribution from the emitted crossover gamma-rays. The observed ratios r_o were calculated from the emission rates and their uncertainty estimates calculated by SAMPO 90 version 3.50. The calculated ratios r_c were taken from the coincidence correction calculation of SHAMAN version 1.13, with the exception of the four sum peaks from non-consecutive gamma-rays in Table 4.7 that were calculated on a spreadsheet. No uncertainty estimates were available for coincidence corrections, because error propagation in the calculation is complicated. However, a crude estimate for the uncertainty of r_c can be obtained with the help of Eqs. 4.20 and 4.21: the area ratio in a simple two-photon cascade is inversely proportional to the peak efficiency, and therefore, its uncertainty must be of the order of the peak efficiency uncertainty. In the test spectra, these uncertainties vary in the range 3 %–6 %.

One could argue that the uncertainty estimate of the observed peak ratio r_o is unnecessarily large, because it contains the contribution of efficiency uncertainty both for the cascade gamma-ray and the sum gamma-ray. These uncertainties are in fact likely to cancel out. However, there are two counter-arguments. First, the contribution of peak area uncertainty is dominant for a majority of the cases since sum peaks are typically quite small, and second, the agreement between r_o and r_c is not so good even with these overestimated uncertainties of r_o. The reasons for the latter are discussed below.

The agreement between the calculated ratios and the observed ones seems to be reasonable. Deviations seem to be of random nature, i.e., some of the calculated sum peak areas are larger than observed and others smaller. This can most readily be seen from the column "Sum peak explanation" where explanation levels of sum peak areas are presented assuming an explanation of exactly 100 % for the corresponding cascade gamma-ray peak. This is equal to $100\,\% \times r_o/r_c$. The explanation levels of the coincidence sum peaks vary typically in the range 70 %–130 %. This is only slightly more than the typical variation of explanations of ordinary photopeaks. This is acceptable as the sum peaks are generally among the smallest peaks in a spectrum.

There are 14 coincidence sum peaks of the total of 55 sum peaks whose explanation level is out of the above mentioned range in Tables 4.7 and 4.8. The reasons for these discrepancies fall into three categories:

A. Interfering photopeak. A significant photopeak component is included in either the sum peak or the cascade peak in the following cases (shares based on activities calculated in a least squares sense by Shaman):

1. 880.0 keV of ^{99}Mo in `stuk3` (Table 4.7): an interference of ^{99m}Tc 140.5 keV line with a peak share 94 %;
2. 1400.5 keV of ^{134}Cs in `stuk7` (Table 4.7): an interference of ^{214}Bi 1401.5 keV line with a peak share 38 %;
3. 334.3 keV of ^{239}Np in `stuk7` (Table 4.8): an interference of ^{132}Te 228.16 keV line with a peak share 51 %;
4. 920.6 keV of ^{99}Mo in `stuk8` (Table 4.7): an interference of ^{140}La 919.55 keV line with a peak share 90 %;
5. 1297.9 keV of ^{132}I in `stuk8` (Table 4.8): an interference of ^{133}I 1298.223 keV line with a peak share 40 %;
6. 1400.5 keV of ^{134}Cs in `stuk8` (Table 4.7): an interference of ^{214}Bi 1401.5 keV line with a peak share 54 %;
7. 334.3 keV of ^{239}Np in `stuk8` (Table 4.8): an interference of ^{132}Te 228.16 keV line with a peak share 43 %;
8. 334.3 keV of ^{239}Np in `stuk9` (Table 4.8): an interference of ^{132}Te 228.16 keV line with a peak share 65 %.

B. Angular correlation. The effect of angular correlation is neglected in the calculation of coincidence corrections. This should give reasonably good results for all other transitions but (0+2+0+)-transitions whose angular correlation is very strong: an angular correlation value $W(0) = 2.3$ is presented in Ref. [105]. This effect is clearly seen as an underexplanation of the two sum peaks corresponding to (0+2+0+)-transitions:

9. 1133.8 keV of ^{106}Rh in `stuk2` (Table 4.7): explanation level 47 %;
10. 1837.3 keV of ^{42}K in `stuk3` (Table 4.7): explanation level 55 %.

C. Unknown explanation. No specific explanation can be found for the clear under- or overexplanations listed below. The major uncertainty contributions in observed peak area ratios r_o stem from area determination and peak efficiency calibration that should be quite well in control. However, areas of small peaks can sometimes be quite inaccurate, especially if they are components in a multiplet. Calculated peak area ratios r_c depend on peak and total efficiencies and decay scheme data. Especially inaccuracies in decay scheme data, often simplified from the full schemes, can lead to large errors, but these errors are systematic and can be revealed by measuring the same nuclide with different spectrometers. However, it should be noted that we are investigating single measurements. In order to evaluate the estimated uncertainties, we would need a statistical analysis of a large number of spectra.

The following peaks belong to this category:

11. 1542.4 keV of ^{110m}Ag in `stuk1` (Table 4.7): explanation level 67 % probably due to decay scheme inaccuracies;

12. 1562.3 keV of ^{110m}Ag in `stuk1` (Table 4.8): explanation level 63 % probably due to decay scheme inaccuracies;

1. 880.0 keV of ^{99}Mo in `stuk3` (Table 4.7): when the interference with ^{99m}Tc is accounted for, the sum peak explanation level is only 50–60 %, probably due to decay scheme inaccuracies;

13. 1440.3 keV of ^{132}I in `stuk7` (Table 4.7): explanation level 136 % probably due to small peak area since the same sum peak is correctly explained in two other spectra;

14. 350.6 keV of ^{143m}Ce in `stuk8` (Table 4.8): explanation level 60 % probably due to small peak area.

Even if a full explanation of a gamma-ray spectrum is the aim, one has to allow a larger deviation from a 100 % explanation level for coincidence sum peaks than for ordinary photopeaks. This is mainly due to the last two items in the list above. However, no substantial differences in the explanations of sum peaks can be seen between Table 4.7 containing pure or almost pure sum peaks and Table 4.8 containing peaks with non-zero initial emission probabilities. This is a demonstration of the fact that coincidence correction calculation always introduces an additional uncertainty to apparent emission probabilities. In other words, nuclide activities can be determined more accurately in a measuring setup where coincidence summing is negligible.

4.7.3 Conclusion

Applicability of the method of Andreev et al. [101] was verified for prediction of coincidence sum peak areas. This method is widely used in the calculation of coincidence corrections for photopeaks, but it can also be applied to coincidence sum peak calculations as was demonstrated above. The modification needed in the basic calculation method is minimal: the zero or low probability crossover transitions are added to decay schemes with a small initial emission probability that is subtracted from the calculated emission probabilities afterwards. This calculation has been implemented in Shaman.

The prediction of sum peaks from non-consecutive transitions can be integrated into the calculation method, but this has not yet been done. More extensive modifications in Shaman are required for implementing calculation of X-ray sum peaks. Automated manipulations of decay schemes to account for annihilation gamma-rays and metastable states is another large development challenge.

The calculated coincidence sum peak areas agree with the observed ones typically within 30 %. This is a sufficient accuracy for nuclide identification, especially when one takes into account that sum peaks often have relatively large area uncertainties due to their smallness.

4.8 X-Ray Escape Peaks

X-ray escape peaks are a result of germanium X-rays escaping from the active volume of the detector, leading to an energy transfer deficit of 10 keV and when abundant, a peak 10 keV below the photopeak. Calculation methods for characteristic X-ray escape peak areas in Ge-detector spectra have been presented by several authors [49, 117–121]. Different aspects of the problem are handled by these authors, including the angular distribution of the incident radiation in Ref. [117], the two groups K_α and K_β of Ge X-rays in Ref. [119], and the dependence of X-ray escape peak areas on incident angle and the irregularities of the detector surface in Ref. [121].

Without an exception, however, these authors have concentrated on the energy region below ∼100 keV where escape through the sides and rear of the detector can be neglected. The motivation is that the escape-peak-to-photopeak ratio, denoted here as X, decreases significantly as a function of energy. Another motivation has probably been that above 100 keV the significance of Compton scattering begins to be proportional to, and eventually higher than, that of photoelectric effect. This makes calculations more complicated as will be seen in the following.

Even above 100 keV, however, the ratio X can be of the order of 10^{-3} and it remains at a relatively constant level throughout the usual energy region of interest in gamma-ray spectrometry. In favorable conditions this means that an X-ray escape peak can be seen in a spectrum much above 100 keV, especially when the continuous improvement of Ge-detector resolution and charge-collection is taken into consideration. As an illustration from airborne radioactivity monitoring, X-ray escape peaks are occasionally seen below the strong 238 keV peak of ^{212}Pb that is present in air as a relatively long-lived descendant of ^{224}Rn.

We present a method for calculating X-ray escape peak areas that extends the applicable energy range to about 2 MeV. Its predictions are compared to empirical data as well as Monte Carlo simulated data in order to verify whether the accuracy of the method is sufficient for gamma-ray spectrum interpretation purposes. Regarding the numerical data needed in the calculations, photon attenuation coefficients have been taken from Ref. [122] and fluorescence yields from Ref. [114].

4.8.1 Calculation Method

In the following, calculation of the X-ray escape probability P is divided into six separate components:

1. escape from the *front* surface following direct photoelectric interaction (P_{pf});
2. escape from the *rear* surface following direct photoelectric interaction (P_{pr});
3. escape from the *side* surface following direct photoelectric interaction (P_{ps});
4. escape from the *front* surface following one or several Compton-scatterings and a final photoelectric interaction (P_{cf});
5. escape from the *rear* surface following one or several Compton-scatterings and a final photoelectric interaction (P_{cr});
6. escape from the *side* surface following one or several Compton-scatterings and a final photoelectric interaction (P_{cs}).

At energies below 100 keV, the first component is predominant, but the components involving Compton-scattering become significant above this energy. This is a direct result of the energy dependence of the respective attenuation coefficients.

Escape from the front surface following direct photoelectric interaction

Assume that the angle θ made by the incident radiation with the detector axis is 0, corresponding to a point source in infinity or a collimated source. Then the escape probability from the front surface of the detector P_{pf} following direct photoelectric absorption can be derived assuming a semi-infinite detector and using the following quantities:

1. The differential probability dP_1 for first photon interaction between x and $x + dx$ is
$$dP_1 = \mu_\gamma \exp(-\mu_\gamma x)\, dx \ ,$$
where μ_γ is the total attenuation coefficient in germanium at the incident gamma-ray energy and x is the depth from the surface of the active Ge-volume.
2. The differential escape probability dP_2 for a Ge X-ray from depth x is[e]
$$dP_2 = \frac{1}{2}\exp(-\mu_x x/\cos\phi)\sin\phi\, d\phi \ ,$$
where μ_x is the total attenuation coefficient at the X-ray energy and ϕ is the angle made by the X-ray with the detector axis (x-axis).
3. The share of K-shell ionizations of all Ge ionizations produced by photoelectric effect is q_K. According to Ref. [118], it is equal to
$$q_K = 1 - \mu(E_{\mathrm{KE}} - \delta)/\mu(E_{\mathrm{KE}} + \delta) \ ,$$
where E_{KE} is the energy of K-absorption edge in germanium (=11.1031 keV [114]) and δ is an infinitesimal energy.
4. By definition, the share of X-ray emissions of all K-shell de-excitations is the K-shell fluorescence yield ω_K (=0.5353 [114]).
5. By definition, the share of photoelectric interactions of all interactions is the ratio $p = \mu_{\gamma p}/\mu_\gamma$ where $\mu_{\gamma p}$ is the photoelectric attenuation coefficient.

The probability P_{pf} of producing a full X-ray escape is obtained by multiplying these five quantities and integrating over the detector thickness H in the x-direction and over the angles from 0 to $\pi/2$ in the ϕ-space:

$$P_{pf} = p q_K \omega_K \int_0^H dx \int_0^{\pi/2} d\phi \left[\frac{1}{2}\mu_\gamma \exp(-(\mu_\gamma + \mu_x/\cos\phi)x)\sin\phi\right] \ . \tag{4.25}$$

---

[e] This can be derived as follows. Consider a point at depth x, where an X-ray is created by photoelectric interaction, and let the X-ray make an angle ϕ ($\phi \in [0,\pi]$) with the normal of the detector surface. Attenuation of the X-ray is $a = \exp(-\mu_x x/\cos\phi)$. An infinitesimal ring-like surface element, for which the attenuation is constant, is $dA' = 2\pi(r\sin\phi)\times r\, d\phi = 2\pi r^2 \sin\phi\, d\phi$, where $r = x/\cos\phi$ is the radial coordinate in the spherical coordinate system. For normalization, dA' must be divided by the integral over the spherical surface, or $dA = dA'/4\pi r^2 = \frac{1}{2}\sin\phi\, d\phi$. The differential escape probability dP_2 is the product of a and dA.

Integration over x yields:

$$P_{pf} = \frac{1}{2} p q_K \omega_K \int_0^{\pi/2} d\phi \, \frac{\mu_\gamma \sin\phi}{\mu_x/\cos\phi + \mu_\gamma} \left[1 - \exp(-(\mu_x/\cos\phi + \mu_\gamma)H)\right] \ .$$

The second term in the brackets can be set to zero in all realistic cases, since $\mu_x \in [14, 20]\,\mathrm{mm}^{-1}$ and $H \gg 1\,\mathrm{mm}$. In other words, any detector is a semi-infinite detector for X-rays and the x-integration can actually be done from zero to infinity. Denoting the attenuation coefficient ratio by $r = \mu_x/\mu_\gamma$ and integrating over the ϕ-values yields:

$$P_{pf} = \frac{1}{2} p q_K \omega_K \int_0^{\pi/2} \frac{\sin\phi \, d\phi}{r/\cos\phi + 1} = \frac{1}{2} p q_K \omega_K \int_0^1 \frac{t}{t+r}\, dt = \frac{1}{2} p q_K \omega_K \left[1 - r \log(1 + 1/r)\right] \ . \tag{4.26}$$

Consider now the case where the incident radiation makes a non-zero angle θ with the detector axis. This modification leads to the following replacement in Eq. 4.26 [117]:

$$r \to r\cos\theta \ \Rightarrow \ P_{pf} = \frac{1}{2} p q_K \omega_K \left[1 - r\cos\theta \log(1 + 1/(r\cos\theta))\right] \ .$$

If the incident beam is uncollimated, radiation will hit the detector with angles in a range $[\theta_1, \theta_2]$. In this case, integration yields the following expression for the escape probability [117]:

$$P_{pf} = \frac{p q_K \omega_K}{4(\cos\theta_2 - \cos\theta_1)} \int_{\cos\theta_1}^{\cos\theta_2} \left[\cos\theta + r\cos^2\theta \log(r\cos\theta) + \left(\frac{1}{r} - r\cos^2\theta\right) \log(r\cos\theta + 1)\right] . \tag{4.27}$$

The formulation of Eq. 4.27 is the final result for an uncollimated radiation source, whereas the simpler form of Eq. 4.26 can be used for a collimated source or when the source-detector distance is large in comparison to the detector radius.

In the case of a pointlike source close to the detector, $\cos\theta_1 = 1$ and the value of θ_2 can be calculated with sufficient accuracy from the detector radius R and the source-detector distance d: $\theta_2 = \arctan(R/d)$. It must be noted that d must be the distance between the source and the *active volume* of the detector, i.e., it must include the distance to the endcap, the thickness of the entrance window, the thickness of the vacuum, and the thickness of the inactive Ge-layer. However, the calculation of P_{pf} is relatively insensitive to the exact value of θ_2: a change from $\theta_2 = 0°$ to $\theta_2 = 60°$ increases the escape probability only with <40 % at typical energies. Only in the shortest distances does the calculation become sensitive to the exact angle. This is illustrated in Fig. 4.12.

For an extended source, an estimate for the average maximum angle θ_2 could be calculated from the geometry, but it can as well be derived empirically by trial and error. For illustration, the value $\theta_2 = 60°$ has provided results with sufficient accuracy for a case where a compressed air filter sample with 20 mm radius and 10 mm thickness, in a cylindrical plastic container, has been measured on top of a Ge-detector of 18 % relative efficiency. According to Ref. [100], this would correspond to an active detector volume of $79\,\mathrm{cm}^3$, or a detector radius $R \approx 23\,\mathrm{mm}$, provided that the length of the detector is $H \approx 2R$.

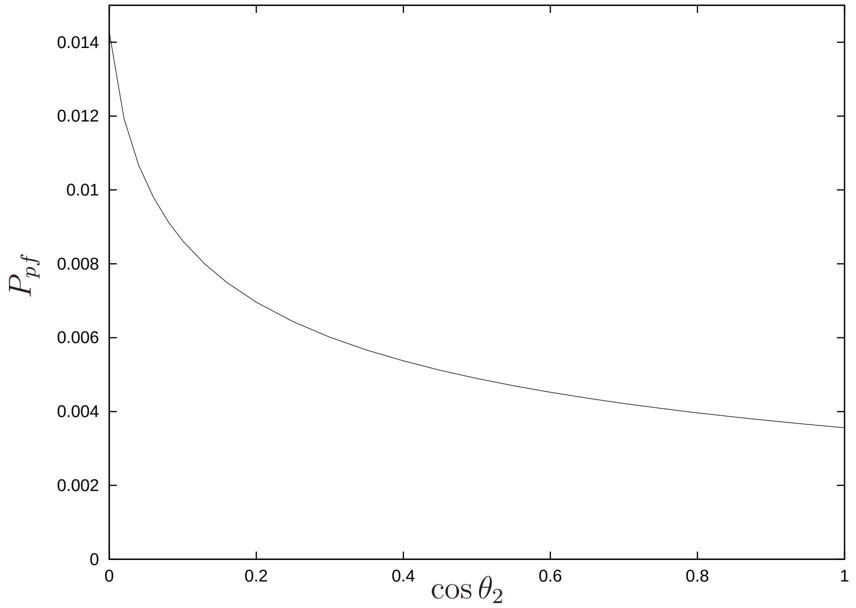

Figure 4.12: Effect of the maximum incident angle θ_2 on the escape probability P_{pf} at photon energy 70 keV.

Escape from the rear surface following direct photoelectric interaction

The same method as for escapes from the front surface can also be used for escapes from the rear surface, when considering the incident photons interacting directly with photoelectric effect in the detector. The probability P_{pr} is obtained from an equation similar to Eq. 4.25. The difference is that the X-rays must escape from the surface $x = H$, and thus, the integration in ϕ-space is from π to $\pi/2$:

$$P_{pr} = pq_K\omega_K \int_0^H dx \int_\pi^{\pi/2} d\phi \left[\frac{1}{2}\mu_\gamma \exp(-\mu_\gamma x - \mu_x(H-x)/\cos\phi)\sin\phi\right] .$$

Denoting again the attenuation coefficient ratio by $r = \mu_x/\mu_\gamma$, we obtain:

$$\begin{aligned} P_{pr} &= \frac{1}{2}pq_K\omega_K \exp(-\mu_\gamma H)\int_\pi^{\pi/2} \frac{\sin\phi\, d\phi}{r/\cos\phi - 1} = \frac{1}{2}pq_K\omega_K \exp(-\mu_\gamma H)\int_{-1}^{0} \frac{t}{t-r}\, dt \\ &= \frac{1}{2}pq_K\omega_K \exp(-\mu_\gamma H)\left[1 - r\log(1+1/r)\right] . \end{aligned} \quad (4.28)$$

This equation is exactly of the same form as Eq. 4.26 for escapes from the front surface, when a narrow incident beam is assumed: $P_{pr} = \exp(-\mu_\gamma H) \times P_{pf}(\cos\theta_2 = 1)$. Due to the exponential factor, this term is practically insignificant in the energy region where photoelectric effect is dominant. Thus, generalization of P_{pr} for other $\cos\theta_2$-values can be omitted and Eq. 4.28 can be used in the calculations.

Escape from the side surface following direct photoelectric interaction

Above the exponential distribution of photoelectrons in the axial direction was utilized for escape probabilities through the front and rear surfaces. For the side surface, the situation is different: we assume that *the distribution of photoelectrons in the radial direction is constant.* This assumption is exactly valid only for a point source in infinity, but it has proven to work with sufficient accuracy even for other geometries.

Let us consider a slab of thickness T, where the distribution of photoelectric absorptions $S(x)$ is constant: $S(x) = 1/T \;\Rightarrow\; \int_0^T S(x)\,dx = 1$. Analogously to Eq. 4.25, the escape probability through one of the slab surfaces is:

$$P'_{ps} = pq_K\omega_K \int_0^T dx \int_0^{\pi/2} d\phi \left[\frac{1}{2T}\exp(-\mu_x x/\cos\phi)\sin\phi\right] .$$

Integration over x yields:

$$P'_{ps} = pq_K\omega_K \int_0^{\pi/2} \frac{\sin\phi\cos\phi\,d\phi}{2\mu_x T}\left[1 - \exp(-\mu_x T/\cos\phi)\right] .$$

The second term in the brackets can again be set to zero in all realistic cases. Integrating over ϕ gives:

$$P'_{ps} = \frac{pq_K\omega_K}{2\mu_x T}\int_0^{\pi/2} \sin\phi\cos\phi\,d\phi = \frac{pq_K\omega_K}{4\mu_x T} .$$

The escape probability through both slab surfaces must be equal, so the total escape probability from the slab is twice the calculated value. When the same calculation is done for a cylinder, an analogous result is obtained: only the dimension T is replaced by R.[f]

In order to obtain the correct normalization, we must also take into account that only the fraction $(1 - \exp(-\mu_\gamma H))$ of the incident gamma-rays interact within the detector volume, assuming a point source in infinity. When the escape probability through side surface is multiplied with these normalization factors, the following result is obtained:

$$P_{ps} = 2 \times \frac{T/2}{R/2} \times (1 - \exp(-\mu_\gamma H)) \times P'_{ps} = \frac{pq_K\omega_K}{2\mu_x R}(1 - \exp(-\mu_\gamma H)) . \tag{4.29}$$

Escape from the front surface following one or several Compton-scatterings and a final photoelectric interaction

At energies above 100 keV, the significance of the processes involving one or several Compton-scatterings and a final photoelectric absorption becomes comparable to and eventually larger than that of the direct photoelectric absorption processes, both for the photopeak and for the corresponding X-ray escape peak. The Compton-processes complicate the calculation due to the multi-step attenuation before the photoelectric absorption and the subsequent emission of the characteristic X-rays. For realistic calculations in this energy region, a Monte Carlo calculation code would be best suited. However, it seems that an acceptable accuracy can be obtained by assuming that *the distribution of photoelectric absorptions following one or several Compton-scatterings is constant in both the axial and the radial direction.*

---

[f] The characteristic dimension d is dictated by the volume-to-area ratio. For a slab, the ratio of the volume to the side surfaces is $d_s = (abT)/(2ab) = T/2$, where a and b are the slab dimensions in the other two directions. For a cylinder, the ratio for the side surfaces is $d_{c1} = (\pi R^2 H)/(2\pi R H) = R/2$, and for the front or rear surface $d_{c2} = (\pi R^2 H)/(\pi R^2) = H$.

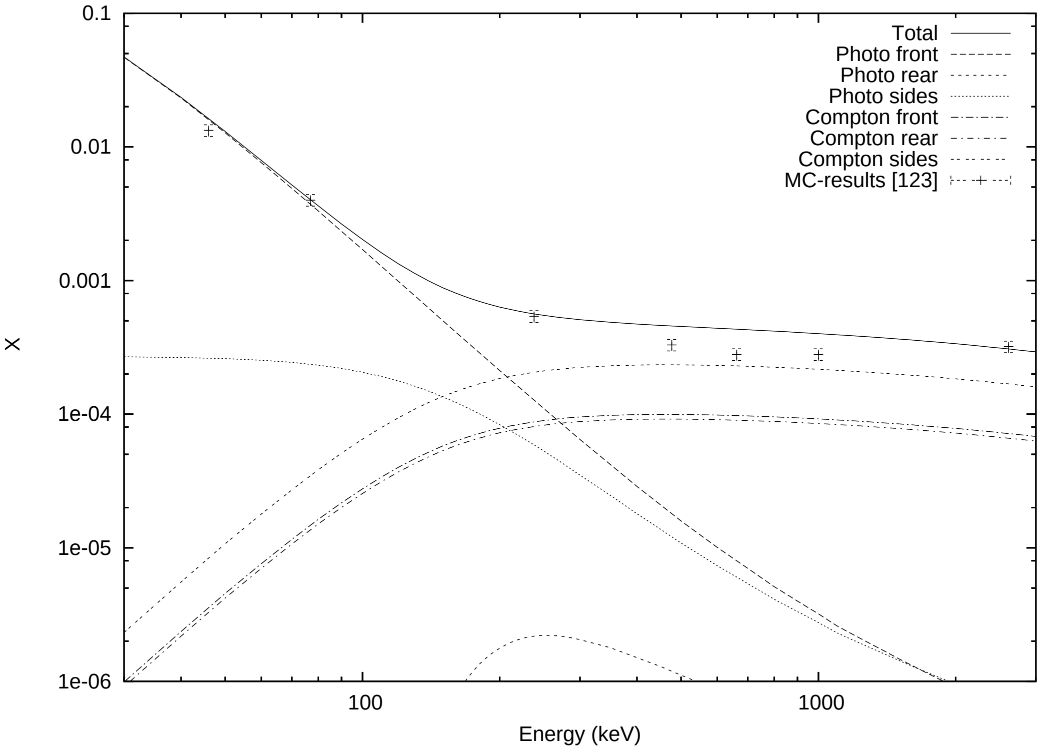

Figure 4.13: The escape-peak-to-photopeak ratio X as a function of energy, when the escape of Compton-scattered photons (last factor in Eqs. 4.30, 4.31, and 4.32) is omitted. The detector has a 100 % relative efficiency with $R = 45\,\mathrm{mm}$ and $H = 55\,\mathrm{mm}$, and the source is a cylindrical source on top of the detector with radius 39 mm and thickness 9 mm.

With this assumption, the equation obtained for P_{ps} above can be applied in the calculation of P_{cf}, the escape probability from the front surface following Compton-processes. As a first approximation we obtain:

$$P'_{cf} = \frac{c}{p} \times \frac{R/2}{H} \times P_{ps} = \frac{cq_K\omega_K}{4\mu_x H}(1 - \exp(-\mu_\gamma H)) \ ,$$

where $c = \mu_{\gamma c}/\mu_\gamma$ or the share of Compton interactions of all interactions, and the factor $(R/2)/H$ comes from volume-to-area ratio considerations.

When comparing the calculated values to Monte Carlo simulation results [123], it appears that this formula overestimates the escape probability (Fig. 4.13). The explanation is that part of the Compton-scattered photons escape from the detector without further interactions and this loss must be compensated for. This can be done by calculating an *averaged energy-dependent non-escape probability for the scattered photon*. As a first approximation, let us assume an average scattering angle $\hat{\alpha}$.[g] The energy of the scattered photon is:

$$\hat{E}'_\gamma = \frac{E_\gamma}{1 + \frac{E_\gamma}{511\,\mathrm{keV}}(1 - \cos\hat{\alpha})} \ .$$

The corresponding attenuation coefficient is $\hat{\mu}_{\gamma'} = \mu(\hat{E}'_\gamma)$. Let us further assume that the

---

[g] $\hat{\alpha}$ can be interpreted as a second adjustable parameter in the model. The first one is θ_2.

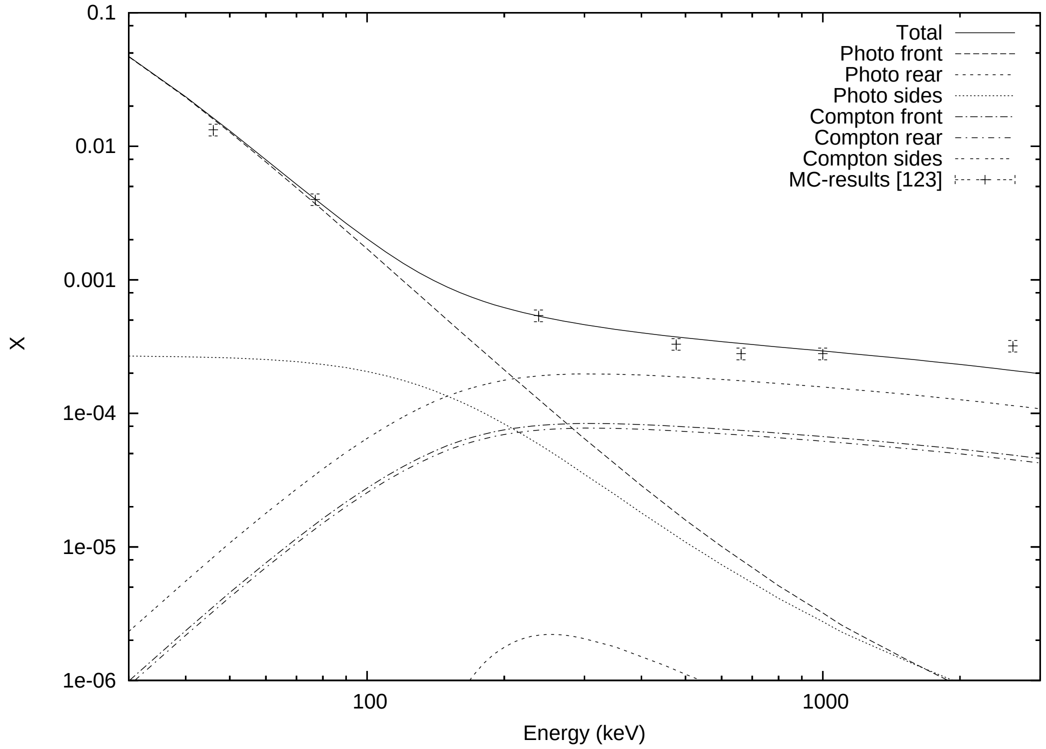

Figure 4.14: The escape-peak-to-photopeak ratio X as a function of energy, when the escape of Compton-scattered photons is taken into account with the simple model presented in the text. Otherwise, the calculation parameters are the same as for Fig. 4.13.

average path length for the scattered photon in the detector is $\hat{x} = HR/(H+R)$.[h] Now we obtain the normalization factor, the non-escape probability for the scattered photon:

$$f = 1 - \exp(-\hat{\mu}_{\gamma'}\hat{x}) \ .$$

With this normalization, the escape probability becomes:

$$P_{cf} = \frac{cq_K\omega_K}{4\mu_x H}(1 - \exp(-\mu_\gamma H))(1 - \exp(-\hat{\mu}_{\gamma'}\hat{x})) \ . \tag{4.30}$$

An excellent agreement with Monte Carlo -calculations is seen in Fig. 4.14. The improvement from Fig. 4.13 between 200 and 1000 keV is obvious.

Escape from the rear and side surfaces following one or several Compton-scatterings and a final photoelectric interaction

With the assumption of constant distribution of photoelectric absorptions following Compton-scatterings in the axial direction, the escape probability from the rear surface is equal to that from the front surface:

$$P_{cr} = P_{cf} \tag{4.31}$$

---

[h] The definition is equivalent to $\hat{x} = 2V/A$. If $H = 2R$, the average path length is $\hat{x} = 2/3H$, which is a realistic value.

Analogously, the escape probability from the side surface must be equal to that from the front surface multiplied by the surface area ratio:

$$P_{cs} = \frac{2H}{R} \times P_{cf} \tag{4.32}$$

4.8.2 Experimental and Simulated Results

The calculation method for the X-ray escape peak areas was verified against the following cases:

1. Measured spectra from an 18 % Ge-detector, showing strong photopeaks of ^{212}Pb at 75, 77, and 238 keV.
2. Measured spectra from a 39 % Ge-detector, with $R = 28.65\,\text{mm}$ and $H = 73.4\,\text{mm}$, showing strong photopeaks of ^{137}Cs and ^{60}Co.
3. Measured spectra from a 55 % Ge-detector, with $R = 40\,\text{mm}$ and $H = 30\,\text{mm}$, showing strong photopeaks of ^{60}Co.
4. Monte Carlo simulated spectra from a 100 % Ge-detector, with $R = 45\,\text{mm}$ and $H = 55\,\text{mm}$, for a number of photopeak energies.

The escape-to-photopeak area ratios X measured from real spectra or calculated by Monte Carlo simulation were compared to those calculated from the sum of Eqs. 4.27, 4.28, 4.29, 4.30, 4.31, and 4.32:

$$\begin{aligned} P &= P_{pf} + P_{pr} + P_{ps} + P_{cf} + P_{cr} + P_{cs} \ , \\ X &= \frac{P}{1-P} \ . \end{aligned}$$

Moreover, the escape peaks corresponding to K_α X-ray and K_β X-ray escapes were calculated separately using the respective attenuation coefficients and normalizing the results with the respective fluorescence yields. The energies and relative abundances of these two Ge-X-ray groups can be calculated from the data in Ref. [114, p. F-44]:

$$\begin{aligned} \mathrm{K}_\alpha &: \ E_{K\alpha} = 9.875\ \text{keV},\ \omega_{K\alpha} = 0.474 \ , \\ \mathrm{K}_\beta &: \ E_{K\beta} = 10.982\ \text{keV},\ \omega_{K\beta} = 0.0613 \ . \end{aligned}$$

The KX-energy difference is 1.1 keV, so with a good energy resolution the K_α and K_β X-ray escape peaks are separable. Fig. 4.15 shows that there is a local minimum between the peak components when the width parameter of the Gaussian is below 0.5 keV in an ideal situation. Current detectors reach this resolution at low energies, but at high energies the doublet merges into an asymmetric singlet.

In practice, there are problems in resolving the two components even at low energies. Measured spectra are discrete, they have statistical channel-to-channel fluctuations and the X-ray

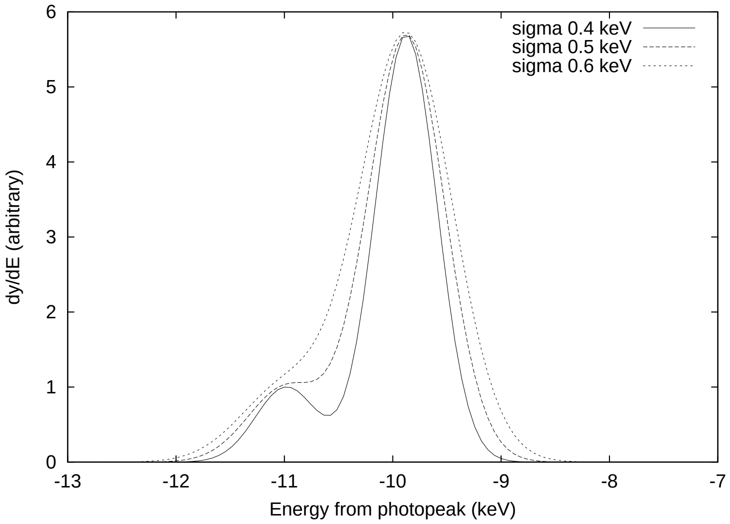

Figure 4.15: The X-ray escape peak components with three different energy resolutions. The figure is theoretical, not measured or simulated.

peaks are in the proximity of a large photopeak. Additionally, the K$_\alpha$ escape peak is approximately a factor of $X_{K\alpha}/X_{K\beta} \approx \omega_{K\alpha}/\omega_{K\beta} \times \mu_{K\beta}/\mu_{K\alpha} \approx 5.7$ times larger than the K$_\beta$ escape peak. This is the ratio that is used in Fig. 4.15. Even if the energy resolution were sufficient, it usually happens that the K$_\beta$ escape peak is below the decision limit and only the K$_\alpha$ escape peak is detected, possibly with some K$_\beta$ contribution on the low-energy side. In Monte Carlo simulations, on the other hand, both escape peaks can be seen when using a sufficiently narrow energy bin and neglecting the finite energy resolution.

In the calculations, the two X-ray escape peaks can be calculated separately. When comparing these results to observations, however, the doublet must be pooled to a singlet, since the doublet is seldom resolved in practice.

Measurements with an 18 % Ge-detector

The first verification of the current method is against spectra from a high-volume aerosol monitoring system (station SE001 of the prototype monitoring network for the CTBT, see p. 98) where a Ge-detector of 18 % relative efficiency was used. The dimensions of this detector are not known to us, but as mentioned previously, they can be estimated with the semi-empirical formula of Ref. [100]: $V \approx 79\,\mathrm{cm}^3$ or $R = H/2 \approx 23\,\mathrm{mm}$. Table 4.9 presents data of the X-ray escape peaks identified in spectra from this detector, measured during the latter half of year 1999. Altogether 129 spectra from this period were available to us and a total of 21 observations of pure X-ray escape peaks were made in 18 of those spectra in pipeline analysis.

The number of X-ray escape peaks in these spectra is not very high. Additionally, their quality is not the best possible, as can be seen from the large emission rate uncertainties and the low significances (peak area relative to the decision limit L_C). Despite these deficiencies, the data in Table 4.9 give a clear indication of the capability of the current method in

Acquisition start date	Peak energy (keV)	Emission rate (gamma-rays/s) with rel. uncertainty	Significance relative to L_C(Shaman)	Predicted peak ratio X	Explanation percentage
1999-09-03	64.88	0.0499 (13 %)	3.26	4.60×10^{-3}	71.43
1999-09-14	64.82	0.1633 (18 %)	2.36	4.60×10^{-3}	136.45
1999-09-15	64.96	0.2616 (11 %)	3.74	4.60×10^{-3}	97.80
1999-11-17	64.59	0.0642 (18 %)	2.16	4.60×10^{-3}	68.91
					Ave. expl.: 94 ± 31
1999-07-09	67.08	0.0334 (19 %)	2.18	4.27×10^{-3}	114.88
1999-07-13	67.32	0.0614 (28 %)	1.34	4.27×10^{-3}	182.09
1999-08-10	67.25	0.1099 (17 %)	2.34	4.27×10^{-3}	122.48
1999-08-13	67.19	0.0150 (34 %)	1.12	4.27×10^{-3}	214.01
1999-08-19	67.25	0.1242 (17 %)	2.44	4.27×10^{-3}	137.77
1999-08-26	67.19	0.1240 (15 %)	2.80	4.27×10^{-3}	104.43
1999-09-03	66.91	0.0385 (16 %)	2.42	4.27×10^{-3}	124.53
1999-09-07	66.78	0.0550 (29 %)	1.30	4.27×10^{-3}	209.76
1999-09-08	67.25	0.0474 (32 %)	1.14	4.27×10^{-3}	220.00
1999-09-10	66.89	0.1156 (10 %)	4.74	4.27×10^{-3}	101.78
1999-09-15	67.06	0.2106 (14 %)	2.90	4.27×10^{-3}	164.81
1999-11-23	67.46	0.0857 (17 %)	2.52	4.27×10^{-3}	87.90
1999-11-24	66.88	0.0462 (26 %)	1.40	4.27×10^{-3}	147.08
					Ave. expl.: 149 ± 46
1999-08-24	228.72	0.0662 (30 %)	1.42	7.12×10^{-4}	99.45
1999-09-07	228.29	0.0702 (24 %)	1.66	7.12×10^{-4}	83.48
1999-09-16	228.83	0.1199 (20 %)	1.92	7.12×10^{-4}	127.38
1999-11-25	228.48	0.0498 (24 %)	1.68	7.12×10^{-4}	54.08
					Ave. expl.: 91 ± 31

Table 4.9: X-ray escape peak explanation levels for an 18 % Ge-detector. The explanation percentage is the ratio of predicted X-ray escape peak area to the observed one in percent. The data are from pipeline analysis of air filter sample spectra with 23 h acquisition time.

explaining X-ray escape peaks: the averaged explanation percentages for the 65, 67, and 228 keV X-ray escape peaks are (94 ± 31) %, (149 ± 46) %, and (91 ± 31) %, respectively. The statistical fluctuation of the explanation percentages is larger than one would expect from the emission rate uncertainties, probably due to inaccuracies in peak fitting in the proximity of a large photopeak. However, these explanation percentages are almost comparable to those of ordinary photopeaks in these and other similar spectra in pipeline analysis when the small peak significances are taken into account.

It is interesting to note that a competing explanation for the 228 keV X-ray escape peak would be the primary photopeak of ^{132}Te. Its presence would indicate a fresh fission product release. By explaining the 228 keV peak correctly as an X-ray escape peak, the number of false alarms can be reduced. It is important to realize, however, that the X-ray escape peak prediction must yield quantitative area estimates in order not to miss a real ^{132}Te event.

Measurements with a 39 % Ge-detector

A couple of spectra were measured with a 39 % p-type Ge-detector using the common calibration sources ^{137}Cs and ^{60}Co [123]. The sources were point sources and they were measured at a distance of 150 mm from the detector inside a lead shield. Relatively high gamma-ray energies were chosen to verify the current method far above the traditional 100 keV limit. The dimensions of the detector, $R = 28.65$ mm and $H = 73.4$ mm, had been determined from an X-ray photograph.

A large discrepancy was found between the X-ray escape peak predictions and the spectra measured with this detector. Very large photopeaks of $> 10^6$ counts were acquired, and according to the calculations, the corresponding X-ray escape peak significances should have been clearly above 5, i.e., easily detectable by a gamma-ray spectrum analysis program. However, no X-ray escape peaks could be detected in these spectra, even if a peak was manually added at the X-ray escape peak energy and fitted.

A possible explanation for this discrepancy is a charge-collection problem in this detector. An incomplete charge-collection is indeed obvious from the relatively poor energy resolution and strong tails in the measured peaks. The effect of this complication on the current method will be elaborated below.

Measurements with a 55 % Ge-detector

Dozens of ^{60}Co spectra measured with a 55 % Ge-detector utilized in a high-volume aerosol monitoring system were available to us. The detector is a large planar p-type crystal of the BEGe-type and its dimensions were kindly provided to us by Mr. R. Werzi of the CTBTO. The detector has a radius of 40 mm and length of 30 mm. The system is operated manually in a daily cycle where a compressed air filter sample is measured for 23.5 hours and a quality control (QC) sample for 15 minutes on the average.

During the first 7 months of 2003, the QC-sample was a relatively active ^{60}Co source that produced spectra with the two characteristic photopeaks at 1173 and 1333 keV, their escape peaks, a coincidence sum peak at 2505 keV and random sum peaks at 2346 and 2665 keV. Additionally, X-ray escape peaks at 1163 and 1323 keV were observed in pipeline analysis, a total of 1+7 observations in a total of 156 QC-spectra. Their data are presented in the upper half of Table 4.10.

The X-ray escape peaks were close to the decision limit. It can be seen that the peaks were underexplained, which means that their sizes were larger than expected. This is probably due to an area overestimation by the fitting algorithm rather than a failure of the escape peak estimation method. This is typical for small peaks in the proximity of a large peak in pipeline analysis. Additionally, testing peak areas against a critical limit inevitably adds a positive bias to the least significant peaks.

The spectrometer was operating very stably during January–July 2003. In order to obtain more accurate X-ray escape peak areas, QC-spectra measured during each month were summed together and the sum spectra were analyzed interactively. This method is not rigorously correct, because the measurements were not adjacent time slices. However, the

Acquisition start date	Peak energy (keV)	Emission rate (gamma-rays/s) with uncertainty	Significance relative to L_C(Shaman)	Predicted peak ratio X	Explanation percentage
2003-05-03	1163.09	19.4666 (25.03 %)	1.70	2.37×10^{-4}	30.34
Ave. expl.: 30 ± x					
2003-01-16	1322.75	9.2924 (31.24 %)	1.36	2.23×10^{-4}	63.94
2003-03-17	1324.04	9.2330 (32.95 %)	1.30	2.23×10^{-4}	66.00
2003-05-06	1323.00	9.1032 (29.93 %)	1.40	2.23×10^{-4}	57.58
2003-05-10	1322.85	11.2058 (29.90 %)	1.82	2.23×10^{-4}	50.46
2003-05-12	1323.52	12.9718 (22.37 %)	1.94	2.23×10^{-4}	42.23
2003-05-17	1323.82	5.3115 (24.62 %)	1.76	2.23×10^{-4}	42.13
2003-06-28	1322.33	4.7678 (27.77 %)	1.46	2.23×10^{-4}	105.99
Ave. expl.: 61 ± 22					
2003-01 (sum)	1163.26	8.6446 (10.45 %)	4.22	2.37×10^{-4}	71.03
2003-02 (sum)	1163.60	5.4053 (17.04 %)	2.36	2.37×10^{-4}	111.68
2003-03 (sum)	1164.27	8.3568 (9.78 %)	3.62	2.37×10^{-4}	71.11
2003-05 (sum)	1163.68	2.8328 (31.49 %)	1.28	2.37×10^{-4}	205.90
2003-07 (sum)	1163.65	11.1087 (10.82 %)	3.62	2.37×10^{-4}	53.57
Ave. expl.: 103 ± 62					
2003-01 (sum)	1322.78	5.6473 (7.74 %)	4.66	2.23×10^{-4}	102.65
2003-02 (sum)	1323.79	5.6797 (8.50 %)	4.20	2.23×10^{-4}	100.57
2003-03 (sum)	1323.45	6.1899 (7.95 %)	4.54	2.23×10^{-4}	90.68
2003-04 (sum)	1324.46	5.5432 (9.54 %)	3.78	2.23×10^{-4}	97.65
2003-05 (sum)	1323.36	6.0361 (7.83 %)	4.58	2.23×10^{-4}	91.41
2003-06 (sum)	1323.23	7.4212 (8.37 %)	4.28	2.23×10^{-4}	74.78
2003-07 (sum)	1322.84	5.0117 (13.04 %)	2.72	2.23×10^{-4}	112.02
Ave. expl.: 96 ± 12					

Table 4.10: X-ray escape peak explanation levels for a 55 % Ge-detector. The explanation percentage is the ratio of predicted X-ray escape peak area to the observed one in percent. The data for the first 8 peaks are from pipeline analysis of individual QC-spectra, the rest from interactive analysis of summed spectra.

peak shapes in summed spectra remained very close to the nominal shape, with a maximum increase from Gaussian width 0.85 keV to 0.89 keV, manifesting the stability of the spectrometer and applicability of spectrum summing.

The X-ray escape peak results for the summed spectra are shown in the lower half of Table 4.10. It is evident that the X-ray escape peaks are more significant in the summed spectra and their explanation level is closer to the ideal 100 %. Especially the 1323 keV escape peaks that have a larger significance produce an average explanation percentage close to the ideal and its uncertainty is consistent with the emission rate uncertainty of these peaks. An example of a 1333 keV peak and the corresponding X-ray escape peak in a summed spectrum is shown in Fig. 4.16. These data prove that X-ray escape peaks can really be observed at energies much above 100 keV and that the current method yields a reasonable estimate for their areas at least for this detector.

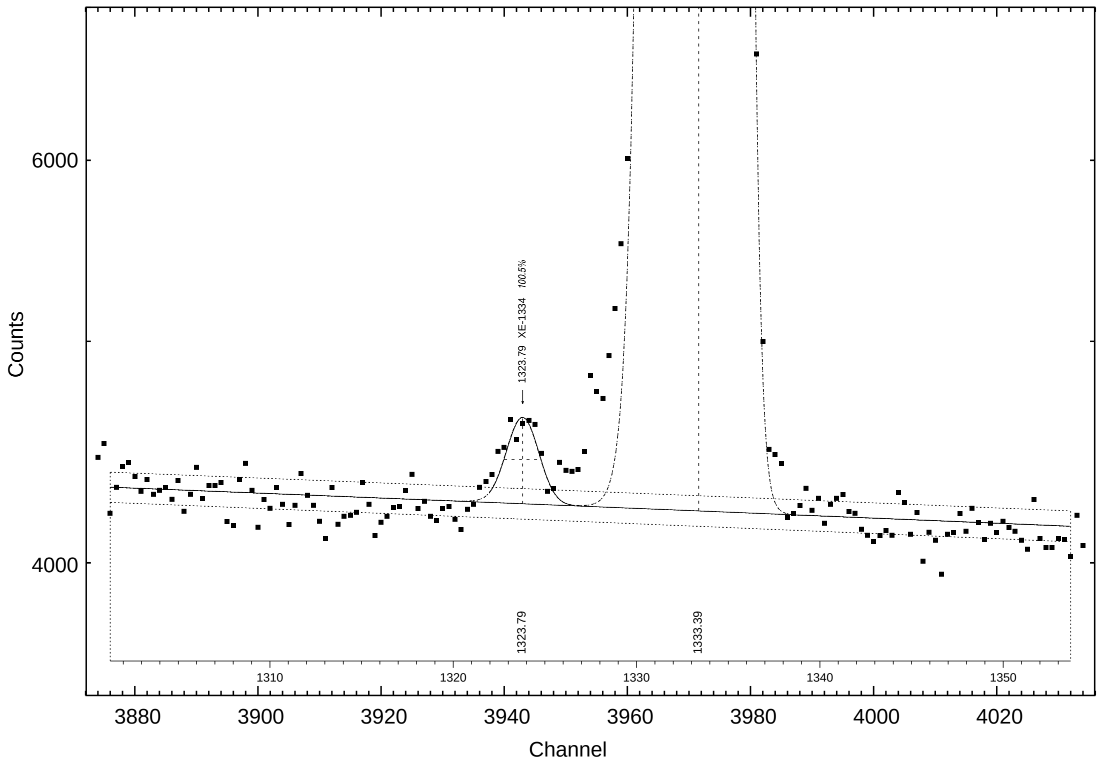

Figure 4.16: An example of a 1333 keV peak with the corresponding X-ray escape peak in a summed spectrum from a 55 % detector. The SAMPO peak model in use is not quite accurate for the lower tail of a photopeak of this size (11 million counts), but the X-ray escape peak is undisputable (2 500 counts).

Monte Carlo simulations for a 100 % Ge-detector

Measurements with a 100 % Ge-detector, with $R = 45$ mm and $H = 55$ mm, were simulated with the MCNP code [124] for a number of gamma-ray energies [123]. The detector geometry, investigated with the help of X-rays, was modeled in details, including the bulletizations and cuts, for various simulation purposes. In the X-ray escape calculations, a 0.2 keV energy bin was utilized and good statistics were obtained by using a large number of gamma-rays: at 1000 keV, for example, 5×10^7 photons were transported.

As also this detector is used in a high-volume aerosol monitoring system, the source used in the calculations was a glass fiber filter compressed into a cylindrical container with $R_s = 38.5$ mm and $H_s = 9$ mm. Calculations with point sources were also performed for comparison purposes, and the X-ray escape results were not very different from those with the cylindrical source.

The results from the Monte Carlo calculations are shown in Table 4.11, using the sum of the two X-ray escape peak areas in the calculation of the escape-peak-to-photopeak ratio X. The results from the current method for the same energies and same detector are also shown in this table. The values of the adjustable parameters in the current method were set to $\theta_2 = \pi/3$ and $\hat{\alpha} = \pi/2$. Since the method is rather insensitive to the values of these parameters, these default values were in fact used in all calculations. The uncertainties in

Peak energy (keV)	Peak ratio X_1 (Monte Carlo)	Peak ratio X_2 (current method)	X_2/X_1
46	1.33×10^{-2}	1.63×10^{-2}	1.23
77	4.0×10^{-3}	4.01×10^{-3}	1.00
238	5.4×10^{-4}	5.35×10^{-4}	0.99
477	3.3×10^{-4}	3.74×10^{-4}	1.13
662	2.8×10^{-4}	3.34×10^{-4}	1.19
1000	2.8×10^{-4}	2.93×10^{-4}	1.05
2614	3.2×10^{-4}	2.09×10^{-4}	0.65

Table 4.11: Comparison of escape-peak-to-photopeak ratios calculated with Monte Carlo simulation and the current method. The sum of the two X-ray escape peak components was used in the calculation.

the MC-simulated peak ratios are of the order of 10 %, so the results of the current method almost agree with the MC-values within uncertainties.

The largest discrepancies are at the low energy end and at the high energy end. The geometry has the largest importance at low energies, so the non-ideal source geometry could provide an explanation for the low-energy discrepancy. At the high energy end, there is no obvious explanation for the discrepancy, but the significance of pair-production interactions, which have been omitted from the current method, starts to be comparable to that of Compton interactions at 2614 keV. This is one possible explanation.

The Monte Carlo simulated peak ratios are plotted in Fig. 4.14 on p. 69, together with those of the current method, as a function of energy. Logarithmic scales often make the agreement look better than it actually is, but in this case, Table 4.11 verifies the good agreement.

4.8.3 Conclusion

A simple model for predicting X-ray escape peak areas in a wide energy range was derived in this work. The results show the significance of the phenomenon also above the traditional cutoff energy 100 keV: the escape-peak-to-photopeak ratio may be of the order of 10^{-4}–10^{-3} far above this energy.

Comparing the predictions with real spectra and spectra created by Monte Carlo simulation shows that *the accuracy of the method is, with certain limitations, acceptable for spectrum interpretation purposes.* However, comparison of the predictions with spectra from a 39 % Ge-detector revealed that the method may fail in some cases.

This failure was addressed to charge-collection problems in the detector. Charge-collection problems are strongest close to detector surfaces, and this is also the region where X-ray escape occurs. The current model actually implies that the transition from active germanium to passive is instantaneous, i.e., there is a sharp boundary. Charge-collection problems on the surface make this transition more gradual, and this will serve to reduce the pure X-ray escape pulses and distribute them into the baseline region between the X-ray escape peak and the photopeak.

There are no mechanisms which could make the X-ray escape peak larger than predicted by the current method, excluding possible contribution from pair production which is excluded from the model. Therefore, it can be stated that the current method yields a prediction of *the maximum X-ray escape peak area,* and the difference between the predicted and observed X-ray peak areas depends on the charge-collection efficiency of the detector. It should be emphasized that charge-collection efficiency is subject to change over detector lifetime, and it is also influenced by the bias voltage applied in each measurement, so one needs to regularly monitor the X-ray escape peak areas. Whether or not prediction of upper limits for escape peak areas is sufficient in gamma-ray spectrum interpretation is application-specific.

One question that has not been addressed yet is the sensitivity of the current method to the accuracy of detector dimensions. The dimensions do not matter in the region below 100 keV, where escape from the front surface following direct photoelectric effect is dominant, but above 100 keV, the detector dimensions are important for the calculation. According to some sensitivity calculations, this method does not necessitate determination of detector dimensions by X-ray photographs, but the dimensions presented by the manufacturer in detector specifications are sufficient. However, determination of detector dimensions from its active volume or relative efficiency, using Ref. [100] and assuming $H = 2R$, can lead to substantial errors above 100 keV, so it cannot be recommended if the dimensions can be determined by any other means.

4.9 Decay Corrections

The activity of an identified nuclide depends on the reference time when the activity is to be quoted, because activities vary with time. After a sample has been collected, the time behavior obeys the simple exponential decay law or a more complex formula, known as Bateman's equations, if the nuclide belongs to a decay chain [125]. Different generalizations of these equations have been published, like the recurrence relation in Ref. [5, p. 51].

During sample collection or generation, on the other hand, the time behavior can be calculated from the following differential equation:

$$dN/dt = -\lambda N(t) + b(t) \ , \tag{4.33}$$

where N is the number of radionuclides of a given kind, λ is its decay constant, and b is the collection or generation rate that is time-dependent in a general case. Decay chains interconnect equations of single radionuclides and so do various nuclear reactions in accelerator and reactor conditions. The equations can be solved numerically and in some special cases analytically.

Natural reference times for activities vary from case to case:

1. The sample has been collected over a *sampling time*, e.g., with an air filter. The reference time can be chosen to be the end or the median of the sampling time. The average activity concentration of a radionuclide during the sampling period can also be calculated if the radionuclide sampling rate as a function of time is known.

2. The sample has been irradiated in a nuclear reactor or an accelerator, and the *irradiation time* is known accurately as well as the particle flux and its time profile. The reference time may be the end of irradiation, but sometimes the saturation activity (activity achieved with an infinite irradiation time) is the desired value.
3. The radioactive sample has been received at a certain time, but its history is unknown. The reference time for activity calculations is usually chosen to be the time when the sample has been received. This is analogous with the end time of sampling/irradiation, when the sampling/irradiation time is unknown and is set to zero.

After the radioactive sample has been created in one way or another, one typically waits for its activity to decay to a lower level. The time parameter connected with the decay before measurement is known as *waiting time* (or cooling time). During the measurement or *counting time* the activity decays further and this decay must also be compensated for in order to obtain correct results. These two time parameters are common to each of the activity measurement situations above.

Let us derive the formulae for the decay corrections in reverse order, i.e., first the counting correction, then the waiting correction, and last the sampling and irradiation corrections. Let the start of sampling or irradiation be time 0, the sampling time t_s, the irradiation time t_i, the waiting time t_w and the counting time t_c. Thus, the measurement starts at time $t_s + t_w$ (case 1 in the list above) or $t_i + t_w$ (case 2) or t_w (case 3) and ends time t_c later. The presentation is restricted to cases where decay chains are in equilibrium, i.e., the radionuclide under investigation decays exponentially obeying its own half-life or that of the longest-lived parent.

4.9.1 Counting Correction

For simplicity, let us derive the counting correction in a case where the radionuclide is a single line emitter and its photopeak is not shared by other nuclides. The formula applies, however, equally well to multiline nuclides whose activities have been determined by least squares fitting of the peak areas. The average activity during counting can be calculated by dividing the area of the photopeak A by the *live* counting time t_l, the peak efficiency ϵ_p at this energy and the emission probability p of the gamma-ray transition:

$$\mathcal{A}_{ave} = \frac{A}{p\epsilon_p t_l} \ .$$

However, the activity of the nuclide decays during the *real* counting time t_c according to the simple exponential decay law and the area of the photopeak comes from the following integral:

$$A = \int_0^{t_c} p\epsilon_p \mathcal{A}_c e^{-\lambda t}\, dt = \frac{p\epsilon_p \mathcal{A}_c}{\lambda}(1 - e^{-\lambda t_c}) \ ,$$

where $\mathcal{A}_c$ is the activity at the start of counting and λ is the decay constant. From these two equations the complete counting correction factor $C_c^0 = \mathcal{A}_c/\mathcal{A}_{ave}$ can be calculated:

$$C_c^0 = \frac{\lambda t_l}{1 - e^{-\lambda t_c}} = L \frac{\lambda t_c}{1 - e^{-\lambda t_c}} \ , \tag{4.34}$$

where $L = t_l/t_c \leq 1$ is the live time correction. Actually, it is conventional for spectrum acquisition equipment to express the live time conversely as a dead time percentage D. The relation between the live time correction and the dead time percentage is $L = 1 - D/100$.

In a general case, the dead time percentage changes over time. If the variation of dead time during the measurement is significant, correction with the average dead time percentage may not be sufficient. Additional corrections are needed especially when the sample contains a mixture of short-lived and long-lived nuclides [49, p. 272].

Since the live time correction can be separated from the actual counting correction, it is customary to define the counting correction factor C_c excluding the live time correction. This definition is used in Sampo and Shaman, for instance, and the live time correction is made separately:

$$C_c = \frac{C_c^0}{L} = \frac{\lambda t_c}{1 - e^{-\lambda t_c}} \ . \tag{4.35}$$

The counting correction tends toward unity, when λt_c tends toward 0, and toward the value λt_c, when λt_c tends toward infinity. Assuming that the uncertainty of λ is σ_λ and that of t_c is σ_{t_c}, an uncertainty estimate for C_c can be calculated as:

$$\frac{\sigma_{C_c}}{C_c} = \sqrt{\left(\frac{\sigma_\lambda}{\lambda}\right)^2 + \left(\frac{\sigma_{t_c}}{t_c}\right)^2} \times \frac{1 - e^{-\lambda t_c} - \lambda t_c \, e^{-\lambda t_c}}{1 - e^{-\lambda t_c}} \ . \tag{4.36}$$

4.9.2 Waiting Correction

During the waiting time the activity decreases according to the simple exponential decay law. Thus, the activity at the start of waiting time $\mathcal{A}_w$ can be calculated from the activity at the start of counting by multiplying with the waiting correction factor $C_w = \mathcal{A}_w/\mathcal{A}_c$:

$$C_w = e^{+\lambda t_w} \ . \tag{4.37}$$

The waiting correction tends toward unity, when λt_w tends toward 0, and toward infinity, when λt_w tends toward infinity. Large values of $\lambda t_w = \ln 2 \, t_w / T_{1/2}$ should, however, not appear in identification results, because they correspond to nuclides with short half-lives that should have been discarded during the analysis. Assuming that the uncertainty of λ is σ_λ and that of t_w is σ_{t_w}, an uncertainty estimate for C_w can be calculated as:

$$\frac{\sigma_{C_w}}{C_w} = \sqrt{\left(\frac{\sigma_\lambda}{\lambda}\right)^2 + \left(\frac{\sigma_{t_w}}{t_w}\right)^2} \times \lambda t_w \ . \tag{4.38}$$

$\mathcal{A}_w$ is often the desired activity value and no further correction factors are needed in such a case.

Sometimes an activity value is requested at a reference time selected more or less randomly. In Sampo and Shaman this kind of a decay correction is referred to as the *predecay correction*. It is analogous to the waiting correction: the equations above give the correction factor and its uncertainty, and the correction is to be applied on a counting corrected activity. The reference time can be either before or after the start of counting, but predecay corrections to a time point before the end of sampling or irradiation, if known, do not make sense.

4.9.3 Sampling Correction

During the sampling time the number of collected radioactive nuclides N obeys Eq. 4.33. If the sampling rate can be assumed constant ($b(t) = b$) and initially there are no radionuclides in the sample ($N(0) = 0$), the following equation can be derived for the activity $\mathcal{A} = \lambda N$:

$$\mathcal{A}(t) = b(1 - e^{-\lambda t}) \; .$$

From this equation we can see that b is equal to the activity with an infinite sampling time and it can be calculated from

$$b = \frac{\mathcal{A}_w}{1 - e^{-\lambda t_s}} \; ,$$

because the activity at the end of sampling is equal to $\mathcal{A}_w$.

The desired activity value is usually the total activity sampled during the sampling time. The total number of nuclides is simply $t_s b$, so the total activity is $\mathcal{A}_s = \lambda t_s b$. Therefore, the sampling correction factor $C_s = \mathcal{A}_s / \mathcal{A}_w$ is

$$C_s = \frac{\lambda t_s}{1 - e^{-\lambda t_s}} \; . \tag{4.39}$$

It can be seen that the sampling correction formula is analogous to the counting correction formula. The sampling correction tends toward unity, when λt_s tends toward 0, and toward the value λt_s, when λt_s tends toward infinity. Assuming that the uncertainty of λ is σ_λ and that of t_s is σ_{t_s}, an uncertainty estimate for C_s can be calculated as:

$$\frac{\sigma_{C_s}}{C_s} = \sqrt{\left(\frac{\sigma_\lambda}{\lambda}\right)^2 + \left(\frac{\sigma_{t_s}}{t_s}\right)^2} \times \frac{1 - e^{-\lambda t_s} - \lambda t_s \, e^{-\lambda t_s}}{1 - e^{-\lambda t_s}} \; . \tag{4.40}$$

In the case of air filter samples the average activity concentration a (Bq/m^3) of the nuclide during the sampling time is often the referred quantity. It can easily be calculated from the sampling corrected activity value:

$$a = \frac{\mathcal{A}_s}{V_{tot}} = \frac{\lambda t_s \mathcal{A}_w}{(1 - e^{-\lambda t_s}) V_{tot}} \; ,$$

where V_{tot} is the total volume of air pumped through the filter during the sampling time. Some elaboration on the activity concentration is given below.

4.9.4 Irradiation Correction

The differential equation for the number of nuclides during irradiation is analogous to the case of sampling, or Eq. 4.33. The only difference is that $b(t)$ is replaced by $\sigma\Phi(t)N_0$, where σ is the average cross section of the generating reaction, Φ is the irradiating flux and N_0 is the number of stable nuclides in the sample. One is usually interested either in the end of irradiation (EOI) activity, in which case no irradiation correction is needed ($\mathcal{A}_w$ is the EOI activity), or in the saturation activity value $\mathcal{A}_i$, i.e., the activity with infinite irradiation time. Assuming a constant flux for simplicity, the equation for activity is

$$\mathcal{A}(t) = \sigma \Phi N_0 (1 - e^{-\lambda t}) \; .$$

The saturation activity is clearly $\mathcal{A}_i = \sigma \Phi N_0$, equal to the reaction rate in the sample. This can be calculated from $\mathcal{A}_w$ by multiplying with the irradiation correction factor $C_i = \mathcal{A}_i / \mathcal{A}_w$:

$$C_i = \frac{1}{1 - e^{-\lambda t_i}} \; . \tag{4.41}$$

The irradiation correction tends toward infinity, when λt_i tends toward 0, and toward unity, when λt_i tends toward infinity. Assuming that the uncertainty of λ is σ_λ and that of t_i is σ_{t_i}, an uncertainty estimate for C_i can be calculated as:

$$\frac{\sigma_{C_i}}{C_i} = \sqrt{\left(\frac{\sigma_\lambda}{\lambda}\right)^2 + \left(\frac{\sigma_{t_i}}{t_i}\right)^2} \times \frac{\lambda t_i \, e^{-\lambda t_i}}{1 - e^{-\lambda t_i}} \; . \tag{4.42}$$

Small values of $\lambda t_i = \ln 2 \, t_i / T_{1/2}$ correspond to long-lived nuclides and thus, large irradiation corrected activity values may appear in some identification results. Some of these large values may indicate a spuriously identified nuclide as presented on p. 83.

4.9.5 Elaboration on Decay Corrections

Total correction factor and its uncertainty

As an illustration of the corrections, the average radionuclide concentration a in air during the sampling time is calculated by multiplying the "raw" activity value $\mathcal{A}_{ave}$, average activity during counting, by counting, waiting, and sampling corrections, and dividing the result with the total sampled air volume V_{tot}:

$$a = \frac{C_c C_w C_s \mathcal{A}_{ave}}{V_{tot}} \; . \tag{4.43}$$

When calculating an uncertainty estimate for radionuclide concentration a, one must take into account that the uncertainty of decay constant λ appears in each of Eqs. 4.36, 4.38, and 4.40. As basic differential calculus shows, these terms must not be added quadratically but linearly. Thus, the uncertainty estimate for activity concentration is to be calculated as:

$$\begin{aligned}
\left(\frac{\sigma_a}{a}\right)^2 &= \left(\frac{\sigma_{\mathcal{A}_{ave}}}{\mathcal{A}_{ave}}\right)^2 + \left(\frac{\sigma_{V_{tot}}}{V_{tot}}\right)^2 \\
&+ \left(\frac{\sigma_{t_c}}{t_c}\right)^2 \times \left(\frac{1 - e^{-\lambda t_c} - \lambda t_c \, e^{-\lambda t_c}}{1 - e^{-\lambda t_c}}\right)^2 \\
&+ \left(\frac{\sigma_{t_w}}{t_w}\right)^2 \times (\lambda t_w)^2 \\
&+ \left(\frac{\sigma_{t_s}}{t_s}\right)^2 \times \left(\frac{1 - e^{-\lambda t_s} - \lambda t_s \, e^{-\lambda t_s}}{1 - e^{-\lambda t_s}}\right)^2 \\
&+ \left(\frac{\sigma_\lambda}{\lambda}\right)^2 \times \left(\frac{1 - e^{-\lambda t_c} - \lambda t_c \, e^{-\lambda t_c}}{1 - e^{-\lambda t_c}} + \lambda t_w + \frac{1 - e^{-\lambda t_s} - \lambda t_s \, e^{-\lambda t_s}}{1 - e^{-\lambda t_s}}\right)^2 \; .
\end{aligned} \tag{4.44}$$

In practice, the uncertainties of the total air volume and the time parameters are often assumed negligible in comparison to uncertainties of average activity and decay constant. If this is the case, only the first and last term on the right hand side are needed in the uncertainty estimate.

By inspection of the definition of $\mathcal{A}_{ave}$ it can be seen that its uncertainty in Eq. 4.44 can be divided further into four components: the uncertainties of photopeak area A, emission probability p, peak efficiency ϵ_p, and live time correction L. If coincidence correction is applied on the emission probabilities, the uncertainty of correction factor C also enters into the equation. Supposing these are independent random variables, the relative uncertainties are to be added quadratically:

$$\left(\frac{\sigma_{\mathcal{A}_{ave}}}{\mathcal{A}_{ave}}\right)^2 = \left(\frac{\sigma_A}{A}\right)^2 + \left(\frac{\sigma_p}{p}\right)^2 + \left(\frac{\sigma_C}{C}\right)^2 + \left(\frac{\sigma_{\epsilon_p}}{\epsilon_p}\right)^2 + \left(\frac{\sigma_L}{L}\right)^2 . \tag{4.45}$$

It should be noted that the calculations above are simplistic in that a single-peak nuclide without interferences with other nuclides is assumed. When these interferences are taken into account, i.e., nuclide activities are calculated by means of a least-squares algorithm, covariances between different quantities complicate the situation. Still, the uncertainties of total air volume, time parameters and decay constant can be added to the calculated activity uncertainty as shown in Eq. 4.44, since they are independent of each other.

Time dependent flow rate

In the derivations of sampling correction C_s and activity concentration a, the sampling rate b was assumed to be constant. This is a very simplistic assumption. The sampling rate is in fact a product of three time-dependent quantities: nuclide concentration in air $n(t)$ (nuclides/m^3), sampler flow rate $f(t)$ (m^3/second), and filter medium efficiency $e(t)$, i.e., the share of nuclides in air pulled through the filter that remain in it. The time dependency of $e(t)$ should be weak unless the conditions are dusty, leading to clogging of the filter. The same is true for the sampler flow rate $f(t)$ that is the quantity that is easiest to measure. The changes in nuclide concentration in air, on the other hand, are the most complicated. They can be estimated only if the count rate as a function of time is recorded or if a transport calculation for a radionuclide plume has been carried out.

However, if the sampling time is kept constant and preferably below 24 hours, a sufficient estimate for airborne radionuclide concentrations is obtained by neglecting the time dependencies in sampling. In practice, the time dependency of flow rate $f(t)$ should always be recorded, as its average value is preferred over any nominal flow rate value. In order to obtain a numerical estimate for the introduced inaccuracy, we can assume that the flow rate decreases exponentially during the sampling time:

$$f(t) = f_0 e^{-\beta t} \quad \Rightarrow \quad f_{ave} = \frac{1 - e^{-\beta t_s}}{\beta t_s}$$

The activity during sampling can be calculated:

$$\mathcal{A}(t) \quad = \quad b_0 \lambda t e^{-\lambda t} \frac{e^{(\lambda-\beta)t} - 1}{(\lambda - \beta)t} \qquad \beta \neq \lambda$$

$$\mathcal{A}(t) \;=\; b_0\lambda t e^{-\lambda t} \qquad\qquad \beta=\lambda$$

The quantity b_0 is the initial value of sampling rate. In this situation, the ratio r of the activity concentration calculated with a constant-flow-rate assumption to that calculated accurately is:

$$r=\frac{a_{approx}}{a_{exact}}=\frac{\lambda t_s}{e^{\lambda t_s}-1}\frac{\beta t_s}{1-e^{-\beta t_s}}\frac{e^{(\lambda-\beta)t_s}-1}{(\lambda-\beta)t_s}$$

A numerical estimate is obtained by assuming a 1-day sampling period, a radionuclide with a half-life of 10.64 h (^{212}Pb) and $\beta = 0.4$. The value of β corresponds to a situation where the flow rate decreases from 100 % to 67 % during 1 day. When these values are inserted to the equation above, we obtain $r = 0.9505$. In other words, the activity concentration needed to obtain a waiting corrected activity of $\mathcal{A}_w$ is underestimated with 5.2 % when the assumption of a constant flow rate is used. In practice, flow rates should not vary as much as in the example and the half-life in use is among the shortest ones of interest in airborne radioactivity surveillance. Thus, for more realistic flow rate variations and for longer-lived radionuclides, the effect of the flow rate approximation is below 5 %.

The effect of assuming a constant flow rate is negligible in comparison to the effects of variation of radionuclide concentrations during sampling time. Especially if a sharp release plume passes the sampling station over a short time interval, the calculated average activity concentrations can be underestimated by orders of magnitude. This is why sampling times are to be shortened in a fresh release situation.

Usage of irradiation corrected activity

In an irradiated sample, the mass m of the stable nuclides producing the identified radionuclide can be calculated from the saturation activity provided that the irradiating flux and the irradiation time are known accurately enough by using the equation $\mathcal{A}_i = \sigma\Phi N_0$. The connection between N_0 and m is $N_0 = N_A m/M$, where N_A is the Avogadro constant and M is the atomic mass (g/mol). We obtain the following equation for m:

$$m=\frac{N_0 M}{N_A}=\frac{\mathcal{A}_i M}{\sigma\Phi N_A}\;.$$

Now, if the calculated mass exceeds the total mass m_{tot} of the sample, it is possible to make the assertion that either the calculated activity is too large, i.e., the peak share of this nuclide is too large, or the nuclide is spuriously identified and should be replaced by another nuclide. The discarding rule can be formulated in terms of activities, e.g., in the following way:

Rule: If the saturation activity of a candidate nuclide exceeds the value $\sigma\Phi N_A m_{tot}/M$, the nuclide is judged improbable.

It should be noted that this formulation is conservative, since samples generally consist of various nuclides, not just the one of interest.

Let us take a numerical example to see the applicability of this rule. Assume that ^{198}Au with $\lambda = \ln 2/2.69517$ d has been identified, when the irradiation time has been $t_i = 2$ min and the neutron flux $\Phi = 1{\cdot}10^{12}\,1/\mathrm{cm}^2\mathrm{s}$. The constant values are $\sigma = 98.65\,\mathrm{b}$, $N_A = 6.022{\cdot}10^{23}\,1/\mathrm{mol}$, and $M = 197\,\mathrm{g/mol}$. If we assume a sample mass $m_{tot} = 10\,\mathrm{mg}$, the maximum saturation activity accepted by the rule above is $\mathcal{A}_i \approx 3.0\,\mathrm{GBq}$. According to Eq. 4.41, this corresponds to an EOI activity $\mathcal{A}_w \approx 1.1\,\mathrm{MBq}$. If the counting and waiting corrected activity of ^{198}Au exceeds this value, the 411.8 keV peak should be explained partly or completely by some other nuclide.

Chapter 5

Expert System SHAMAN

5.1 Expert Systems in General

An expert system can be defined as a computer program that uses the knowledge of an expert to solve a well-defined problem [126–128]. Ideally it implements a model of the knowledge and inference strategies of skilled human experts in the field. When developing an expert system the major challenge is to represent the knowledge and the inference strategies in a format suitable for a computer.

The knowledge of an expert can usually be described as objects, their classifications and relationships, and rules governing the processing of the knowledge in a specific problem. The rules typically contain heuristic methods, i.e., they are often some kind of rules of thumb or educated guesses that have been accumulated to the human expert over the years. Solutions to problems requiring significant expertise cannot usually be expressed as a strict algorithm, because the inference proceeds in a different way when the input data vary. Therefore, the "traditional" procedural programming is not readily applicable to these problems.

Development of an expert system is feasible when the problem to be solved can be clearly restricted to a specific domain. It shall be possible to express the knowledge of a human expert as detailed rules governing the inference, since a computer is unable to create a rule base from some general axioms, and additionally, it does not have any common sense. The rule base shall be sufficiently comprehensive to enable the computer to mimic the problem solving of the human expert. On the other hand, the problem at hand must not be too easy, because the effort put into the development of an expert system shall be paid back in the savings of expert time usage. A suitable problem is one that takes a human expert more than a few minutes but less than a few months to solve [126]. Typical expert system applications belong to some of the following categories: prediction, analysis/diagnosis, planning, monitoring or instruction.

Typically an expert system is developed by a knowledge engineer who consults a human expert for the details of the problem. The main effort of the knowledge engineer is to conceptualize and formalize the problem to a structure that can be implemented as a computer program. The principal structures to be implemented are *data objects* and their relationships, *rules* needed in the solution and an *inference engine* that governs firing of the rules in an appropriate order. There are commercial expert system development tools that can be used to implement an expert system, but performance or distribution issues may demand a stand-alone version to be developed in a later phase.

5.2 Shaman — an Expert System for Radionuclide Identification

Nuclide identification in gamma-ray spectrum analysis as described in Sec. 4.2 is a task that requires significant expertise, especially if a complete explanation of the spectrum peaks is aimed at. When the peaks found in a gamma-ray spectrum are associated with the gamma-ray lines in a reference library, dozens of nuclides in the library may be considered as candidates for each peak. Unfitting candidate nuclides are discarded one by one according to well-defined pruning rules. The final result should be the nuclide combination that gives the best explanation to the peaks observed in the spectrum.

In nuclide identification, partially heuristic pruning rules are combined with numerical computation tasks and managing of a large reference library. It is difficult to formulate the identification procedure as an algorithm, because it varies significantly from one case to another. However, implementations of procedural nuclide identification have been published in Refs. [10] – [16] and Refs. [52] – [84]. The published results show that relatively simple gamma-ray spectra can be correctly identified with this kind of identification programs provided that a tailored application-specific gamma-ray library is used. If the number of spectrum peaks or the size of the reference library is increased, the performance of these programs begins to decline.

In late 1987, the development of an expert system based nuclide identification program was initiated at the Nuclear Engineering Laboratory (currently Laboratory of Advanced Energy Systems) of Helsinki University of Technology [1]. The requirements for an expert system are clearly fulfilled by the nuclide identification task. It is a well-circumscribed problem and its solution can be formulated as explicit rules. Nuclide identification is also a suitably laborious task: a full analysis of a typical gamma-ray spectrum may take a human expert a couple of hours, the most difficult cases even a few days. Furthermore, considerable expertise on gamma-ray spectrometry had been accumulated in the laboratory during the development of the Sampo software family [8–16].

The development of the expert system Shaman was started within a Lisp-based expert system development tool called GoldWorks [129]. In a later phase, the development tool was abandoned due to performance and portability issues and the expert system was rewritten in ANSI-conformable C-language [130]. Actually, the kernel of Shaman has been written in a dedicated *SHC* rule language, developed by T. Hakulinen [5], that is machine-translated to ANSI-C before the actual compilation. The usage of the *SHC* language eases program maintenance as the *SHC* compiler automatically keeps track of all interrelations in the source files. It creates the necessary header files and some basic handling functions for the data structures. It also enforces internal consistency of the knowledge base when rules and objects are modified.

The implementation of the inference engine of Shaman is conceptually quite simple, supporting only forward chaining as the inference strategy [5]. All rules that have been found to match are put in an agenda, ordered according to their priorities. When a rule is fired, each rule in the agenda is checked, because the firing of one rule may affect the status of the others, eventually disabling or enabling them. These actions keep the inference going. When all enabled rules have been fired, the final solution has been achieved. On current hardware, running a typical gamma-ray spectrum through Shaman typically takes 5–20 seconds of CPU-time.

The expert system Shaman is known to compile and run in several computing environments. The C-language version of Shaman was originally developed under OS/2 (Intel-PC), then under AIX (IBM RS-6000), and since 1995 the development platform has been Linux on an Intel-PC. Over the years, versions for the operating systems Digital Unix, HP/UX, Solaris, and MS-Windows have been built in addition to the development platforms mentioned above. This has made Shaman easily portable to different environments. A detailed technical documentation of Shaman can be found in Ref. [5] and a higher level documentation with operation instructions meant for the end-user in a comprehensive manual, Ref. [XX]. The rule base and methods of Shaman, as well as its user interface, are still under constant development.

5.3 SHAMAN's Rules and Parameters

The heuristic knowledge of an expert doing the nuclide identification task during gamma-ray spectrum analysis has been formulated as a set of abstract rules in Shaman. The current number of abstract rules is 29 as shown in Table 5.1 and most of them are quite easy to understand. The abstract rules have been implemented in Shaman as about 60 physical rules, i.e., many of the abstract rules are implemented as two or several physical ones. This is because it is essential to get the physical rules in a proper sequence for performance reasons. The physical rules are classified into 17 groups that can be controlled separately. Both abstract and physical rules of Shaman are presented in detail in Ref. [XX, Ch. 8], so the list of rules in Table 5.1 and the example in Fig. 5.1 should suffice here.

Shaman rules can be roughly categorized into early and late phase rules. The early phase covers the initial reading phase of candidate radionuclides into the knowledge base. Early phase rules are used for discarding the most improbable candidate nuclides, leading to as small peak-nuclide interference matrix as possible for the first activity calculation. The late phase rules cover the duration of the inference process of Shaman including activity calculations and deductions based on its results. Most of these rules depend on the particular interference setup that is changed after each recalculation of activities. In a typical case, these rules have to be applied multiple times as interference groups change. In contrast to early phase discardings, candidates discarded with late phase rules may be reconsidered by Shaman. However, a reconsidered candidate must not be discarded again with the same rule, because this would lead to infinite loops.

Shaman reports only the major steps of its inference process by default, but a complete inference log is available upon request. The log easily becomes quite long and is usually

1. If any of the nuclide's gamma lines is present in the spectrum, the nuclide is read in as a candidate.
2. No candidates are read in based only on the annihilation peak.
3. No candidates are read in for peaks which Shaman explains as a background, escape, or random sum peak, or which Shaman judges to be insignificant.
4. If the relative significances of the primary and threshold lines of the candidate are close to each other, the candidate is judged improbable.
5. If the only found line of a candidate is an X-ray line, the candidate is judged improbable.
6. If the primary line of the candidate is not found in the spectrum, the candidate is judged impossible.
7. If the absolute significance of the primary line of the candidate is too low at the primary line activity level, the candidate is judged impossible.
8. If neither the candidate nuclide nor any of its parent nuclides have suitable genesis modes, the candidate is judged impossible.
9. If the decay corrected pure nuclide activity of a candidate in the sample is smaller than the smallest detectable activity, the candidate is judged impossible. This does not apply to background nuclides.
10. If the decay corrected activation limiting activity of a candidate in the sample is smaller than the smallest detectable activity, the candidate is judged impossible. This does not apply to background nuclides.
11. If the threshold line is too significant at the theoretical minimum activity of the nuclide, the candidate is judged impossible.
12. If none of the ordinary gamma lines of the candidate is significant at the annihilation line activity level, i.e., when the annihilation line fully explains the annihilation peak, the candidate is judged impossible.
13. If the sample is not purely gaseous and it is contained in an open space, all noble gas candidate nuclides are judged improbable.
14. If the sample is an air filter sample, the background statuses of ^{220}Rn and ^{222}Rn daughters must be selected in accordance with the age of the sample.
15. If there is a candidate whose gamma line emission probabilities match closely (within a given tolerance) to the emission rates of the respective spectrum peaks, all other candidates whose most significant lines are associated to these peaks are judged improbable.
16. If a candidate is associated with a large number of spectrum peaks, it is judged certainly present and is not discarded by any discarding rule.
17. If the candidate nuclide is an isomer whose longer lived parent has not been found, the candidate is judged improbable.
18. If no candidates have been associated to a peak, discarded candidates that have earlier been associated to this peak are reconsidered.
19. If the user has especially told Shaman not to accept the candidate for any reason, it is discarded.
20. If the user has especially told Shaman to accept the candidate for any reason, and the candidate can be associated with at least one spectrum peak, it is not discarded by any discrading rule.
21. If the effective half-life of a candidate is incompatible with the user-given half-life limits, the candidate is judged impossible.
22. If the element or mass number of a candidate is incompatible with the user-given element and mass number limits, the candidate is judged impossible.
23. If the calculated activity of a candidate exceeds the threshold activity, i.e., the theoretical maximum activity of the nuclide, the candidate is judged improbable.
24. If the absolute significance of the primary line of the candidate is too low at the calculated activity level, the candidate is judged improbable.
25. If the calculated activity of the nuclide is negative, the candidate is judged improbable.
26. If the uncertainty of the calculated activity is large, the candidate is judged improbable.
27. If the same peaks are associated to both a parent nuclide and its daughter, which is a shorter lived isomeric state of the parent, the daughter nuclide is judged insignificant.
28. If the same peaks are associated to both a parent nuclide and its non-isomeric longer lived daughter, the parent nuclide is judged insignificant.
29. If a peak is poorly associated, i.e., its relative emission rate residual is large, discarded candidates that have earlier been associated to this peak are reconsidered.

Table 5.1: The list of abstract rules implemented in Shaman [XX].

Rule: *If the calculated activity of the nuclide is negative, the candidate is judged improbable.*

```
RULE NegativeActivity
{
    description "Nuclide ended up with a negative activity in activity calculation";
    priority    0;
    trigger     LEVEL;
    depends     {
                    <NUCLIDE_ITEM *>   $&nuclide_pool.nuc_list,
                    <INFCOND_OBJECT *> &activities_done
                };
    rulegroup   PostActivityRuleGroup;

    IF
    {
        RET(DO->no->checked &&
            DO->no->calculated_activity < 0.0F &&
            IsCondition(&activities_done) &&
            !ExistAssertion(&reconsider,DO->no,WILD,negative_activity,NULL));
    }

    THEN
    {
        Imsg(2,"%s: %s\n",status.name,DO->no->nid);
        MakeAssertion(&unfitting,DO->no,negative_activity,NULL);
    }
}
```

Figure 5.1: Example of Shaman rule. First the abstract rule and then the corresponding physical rule in the dedicated *SHC* rule language of Shaman.

applied only to debugging purposes. For the end-user of Shaman the list of discarded candidate nuclides with the reasoning behind discarding, available both in the graphical user interface and the full identification report, is almost always sufficient.

The inference of Shaman can be controlled on a very deep level by switching off one or several of the 17 rule groups, but this is not very common. Many of the rules in Shaman's knowledge base depend on configurable parameters that offer another, more common way to control Shaman. The default set of parameters has not been optimized for any particular purpose, so there is potential for performance gains with some parameter adjustments. The meaning of many of the configurable parameters is quite evident, but some parameters are applied by several rules and their adjustments may lead to unexpected results. The importance of a sufficiently large test set during parameter optimization cannot be overemphasized.

Since the number of Shaman's parameters is above 60 and since they are documented in Ref. [XX, Ch. 9], a couple of examples suffice here. The single most important Shaman parameter is `sigma_tolerance`, a numerical parameter which indicates the energy tolerance in matching spectrum peaks and the corresponding library lines. Inherently, the necessary tolerance depends on the resolving power of the spectrometer, peak analysis routines, and the quality of energy calibration. In Shaman, the tolerance parameter is expressed in units of the width parameter σ of the Gaussian peak, i.e., the tolerance grows toward higher energies when expressed in keV. For a peak at energy E and width parameter σ, gamma-ray lines in the reference library with energies in the interval $[E - t\sigma, E + t\sigma]$ are associated with the peak by Shaman. Here t is the energy tolerance parameter. In analysis pipelines utilizing Shaman, the value $t = 1.2$ has been found to be close to optimum. If the energy calibration can be made more accurate, the tolerance can be lowered but not much below 0.8, since the position uncertainties of small peaks may be of this order, even if the energy calibration is accurate [IV]. If the energy calibration has poor accuracy, the value of t may need to be raised to 2 or even above, leading to large interference groups and spurious identifications.

Another essential parameter related with peak-line associations is `sigma_vicinity`. It indicates the energy tolerance in considering if a nuclide line is close to a peak, possibly as an unresolved component under its tail, but not quite close enough to warrant a full match. In order to be on the conservative side, no conclusions of a nuclide's presence should be drawn on the basis of this kind of lines. The parameter is also expressed in units of the peak width parameter σ. If the value of `sigma_vicinity` is v and the other symbols like above, then lines with energies in the interval $[E - t\sigma, E + t\sigma]$ are considered *found*, and lines with energies in the intervals $[E - v\sigma, E - t\sigma)$ and $(E + t\sigma, E + v\sigma]$ are considered *found close*. If a nuclide has a primary line with the latter status, its discarding is decided by Shaman on the basis of the secondary line. In pipeline analysis, it sometimes happens that multiplets remain unresolved and therefore, Shaman currently uses a value as large as $v = 7.0$. In careful interactive analysis this parameter should be given a much smaller value.

5.4 Shaman's Reference Library

When it comes to the nuclear data required in nuclide identification, the expert system Shaman relies completely on its nuclide library. Therefore, if the library is of low quality, so will the identification results be. The inverse is not necessarily true, but a high quality library lets one concentrate on developing the expert system itself instead of searching for errors in the reference data.

The original nuclide library of Shaman was based on the compilation of Erdtmann and Soyka [88, 89]. It dates from 1979, but most of its data have been measured in the early 1970's and, therefore, it is already inaccurate and even partly erroneous at current time.

Due to the development of semiconductor detectors and measuring electronics, new, more accurate results on nuclear data become available all the time. The research on nuclear data has been organized by the International Nuclear Structure and Decay Data Network, which comprises several research laboratories throughout the world. The evaluated data are published in two journals: Nuclear Data Sheets and Nuclear Physics. They have been compiled into different databases (e.g., ENSDF, NUDAT), which can be accessed at the National Nuclear Data Center (NNDC) in the USA, at the OECD Nuclear Energy Agency Data Bank (NEADB) in France, or at the IAEA Nuclear Data Section in Austria.

The reference library 2.0.0 of Shaman, created in 1993, is mainly based on these databases. The data were retrieved from the NNDC at Brookhaven National Laboratory (BNL) in

Source	Contents
NUDAT-2002 [23]	Element & mass numbers, half-lives, decay modes etc.
NUDAT-2002 [23]	Neutron capture cross sections and resonance integrals
NUDAT-2002 [23]	Decay photon energies and emission probabilities
ToRI-1999 [132]	Decay photon energies and emission probabilities (Table 5.3)
Lederer-Shirley [112]	Percentage branchings to isomer states
England-Rider [42]	Thermal fission yields of ^{235}U
Erdtmann-Soyka [88, 89]	Genesis modes

Table 5.2: The sources of SHAMAN reference library 3.1.1 data. The majority of data is from NUDAT database.

1992 [113, 131]. Some complementary data were collected from other sources. All data were converted into a suitable binary format for SHAMAN, and the quality of the data was carefully checked [I]. However, no comprehensive verification of the data was conducted due to its vast amount. Almost none of the Erdtmann-Soyka data of the original reference library remained in library 2.0.0.

The reference library 3.1.1 of SHAMAN, created in 2003, is mainly based on the NUDAT database. The data were retrieved from the NNDC in 2002 in the PCNUDAT format that includes an MS-Windows program for data retrievals and the database itself [23]. Some gamma-ray data were also taken from the Table of Radioactive Isotopes (ToRI-1999) [132] that is an ENSDF-based database of the Lawrence Berkeley Laboratory (LBL) and Lund University. Some complementary data were again collected from other sources. The conversion process in 2002–2003 was similar to that in 1992–1993, and the libraries 2.0.0 and 3.1.1 have the same binary structure [7]. The library to be used by SHAMAN can currently be selected by the user, but the newer version 3.1.1 is recommended as it is based on more up-to-date data compilations.

5.4.1 Sources of Nuclear Data in Library 3.1.1

Various data of the radionuclides can be utilized in the identification process: half-lives, decay chains, genesis data, and of course, gamma-ray energies and emission probabilities. Most of these data can be found from the databases NUDAT, ENSDF and ToRI [23, 113, 131, 132]. The required data, which either do not exist in these databases or would have been relatively difficult to extract from them, were fission yield data, genesis mode data, and decay percentages to different isomer states. These data were taken to library 3.1.1 from the publications of England-Rider [42], Erdtmann-Soyka [88, 89] and Lederer-Shirley [112], respectively. The different sources are shown in Table 5.2.

Percentage branchings to different isomer states are not directly available from the NUDAT and ENSDF databases. Hence, they were compiled manually from Lederer-Shirley [112], which is certainly not in full accordance with the NUDAT data due to their age difference. A subject for future development is a program to extract isomer branchings from the ENSDF.

The most important genesis modes of the nuclides are contained in SHAMAN's library. They are based on the compilation of Erdtmann-Soyka, where genesis modes have been divided

^{7}Be	^{11}Be	^{33}Na	^{33}Si	^{34}Si	^{37}P	^{41}Ar	^{58}Cr	^{59}Fe	^{53m}Co
^{81}Ge	^{81m}Ge	^{83}Ge	^{91}Br	^{94}Kr	^{99}Rb	^{80}Y	^{102a}Y	^{102b}Y	^{84}Zr
^{95}Zr	^{103}Zr	^{104}Zr	^{86}Nb	^{95}Nb	^{95m}Nb	^{103}Nb	^{105}Nb	^{88}Mo	^{91m}Tc
^{91b}Ru	^{114a}Rh	^{114b}Rh	^{115}Rh	^{117}Rh	^{94}Pd	^{117}Pd	^{100m}Ag	^{103m}Ag	^{99}Cd
^{124}Cd	^{120c}In	^{122c}In	^{125m}In	^{126m}In	^{130}In	^{130m}In	^{130n}In	^{125}Sn	^{125m}Sn
^{129}Te	^{129m}Te	^{116}Xe	^{117}Xe	^{118}Xe	^{131m}Xe	^{133}Xe	^{133m}Xe	^{135}Xe	^{135m}Xe
^{117}Cs	^{146}Cs	^{148}Cs	^{126}Ba	^{131}Ce	^{135}Pr	^{133}Nd	^{133m}Nd	^{138a}Pm	^{138b}Pm
^{155}Pm	^{137}Sm	^{158}Sm	^{152}Eu	^{160}Eu	^{163}Gd	^{147}Ho	^{153m}Ho	^{171}Ho	^{152}Tm
^{156}Tm	^{157}Tm	^{151}Yb	^{151m}Yb	^{178}Yb	^{162}Lu	^{163}Lu	^{164}Lu	^{168}Lu	^{162}Hf
^{163}Hf	^{164}Hf	^{165}Hf	^{168}Hf	^{171}Hf	^{162}Ta	^{163}Ta	^{169}Ta	^{171}Ta	^{171}W
^{172}W	^{173}W	^{174}W	^{188}W	^{189}W	^{173}Re	^{174}Re	^{192}Re	^{173}Os	^{174}Os
^{176}Os	^{178}Os	^{179}Os	^{196}Os	^{178}Ir	^{182}Ir	^{198}Ir	^{182}Pt	^{185m}Pt	^{180}Au
^{181}Au	^{182}Au	^{184}Au	^{203}Au	^{181}Hg	^{182}Hg	^{187b}Hg	^{189}Hg	^{191b}Hg	^{184}Tl
^{186}Tl	^{188a}Tl	^{188b}Tl	^{189}Tl	^{189m}Tl	^{194b}Tl	^{208}Tl	^{187}Pb	^{187m}Pb	^{195}Pb
^{195m}Pb	^{212}Pb	^{194a}Bi	^{194b}Bi	^{198}Bi	^{212}Bi	^{215}Bi	^{199}Po	^{200}Po	^{202}Po
^{214}Po	^{205}Rn	^{206}Rn	^{224}Rn	^{227}Rn	^{206}Fr	^{232}Ra	^{238}Pa	^{229}U	^{235}U
^{237}U	^{239}Np	^{253}Cf	^{255}Es	^{256}Md	^{257}Rf	^{258}Db			

Table 5.3: The 167 nuclides in library 3.1.1 whose gamma-ray and X-ray data are from ToRI [132].

in a very crude way to the following six classes: genesis by thermal neutron capture (`NTH`), genesis by fast neutron reactions (`NFA`), genesis as a fission product (`NFI`), naturally occurring nuclide (`NAT`), genesis by photon reactions (`PHO`), and genesis by charged particle reactions (`CHA`). Especially the last one includes a vast variety of reactions, and as such, these genesis data are not comprehensive enough for very sophisticated conclusions. Moreover, many nuclides in SHAMAN's current library do not have a counterpart in Erdtmann-Soyka, so they are currently left without any genesis mode information. This is also a subject for future development.

The majority of data in the library 3.1.1 were extracted from NUDAT-2002 [23]. In addition to basic nuclear data and neutron cross sections, it gives gamma-ray line data in a convenient format. However, some gamma-ray line data were either missing from or inaccurate in the NUDAT-2002 database, so the alternative data source ToRI-1999 was used to complement the gamma-ray data in SHAMAN's library. Table 5.3 lists the nuclides whose gamma-ray and X-ray data in the library 3.1.1 are from ToRI.

Additionally, two minor modifications were made to the NUDAT and ToRI data, using the library patching tool developed for this kind of solitary corrections (hence the patch number 1 in 3.1.1):

- The emission probability of the 415 keV line of ^{212}Pb was replaced with (0.001 ± 0.001) in accordance with Ref. [111].
- The energy of the 1620 keV line of ^{212}Bi was replaced with (1620.73 ± 0.23) keV in accordance with Ref. [133].

The number of nuclides in the new SHAMAN library is 3 648, including all 238 stable isotopes, and the number of gamma-ray and X-ray lines is 80 062. All known deficiencies in the old library 2.0.0 were remedied in the new one. In addition to creating a more reliable, up-to-

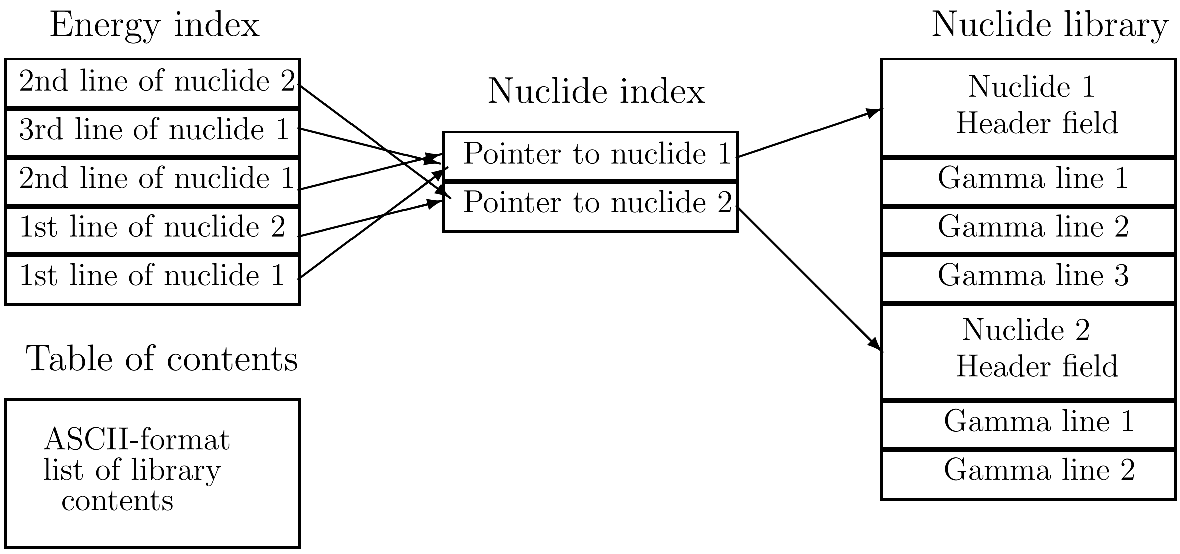

Figure 5.2: The organization of the reference library of Shaman [5].

date library for Shaman, the library update project contributed to correcting several errors in the NUDAT database maintained by the Brookhaven National Laboratory [134].

5.4.2 The Contents of the Library

The nuclide library of Shaman includes the following data concerning well-established radionuclides and their long lived isomer states (and stable nuclides in library 3.1.1):

- element number, mass number and isomer state symbol,
- half-life and its uncertainty,
- direct daughters with percentage branchings,
- direct parents,
- genesis data: fission yield, natural abundance, thermal neutron capture cross section and resonance integral of the generating nuclide, important genesis modes,
- number of gamma-ray lines,
- some status data concerning half-life, fission yield, and decay,
- data of each gamma-ray line of the nuclide:
  - gamma-ray energy and its uncertainty,
  - gamma-ray emission probability and its uncertainty,
  - some status data concerning line type and emission probability.

The reference library file, as well as the nuclide and energy index files, are in a binary format for fast data access. Their organization is illustrated in Fig. 5.2. Two index files are required to enable nuclide searches based on its nuclide key[a] or on its gamma-ray energy. When a nuclide is requested by the expert system, it is searched from the nuclide index file, where the absolute byte address of the nuclide header in the library file is stated. The nuclide header

---

[a] Shaman's internal library search key `NUCLIDE_ID` is calculated from the element number Z, the mass number M and the isomer symbol I using the following formula: `NUCLIDE_ID` $= 2^{16} \times Z + 10 \times M + \mathrm{num}(I)$, where $\mathrm{num}(I)$ = 0,1,2,3,4,5,6,7, when I = G,M,N,O,A,B,C,D, respectively.

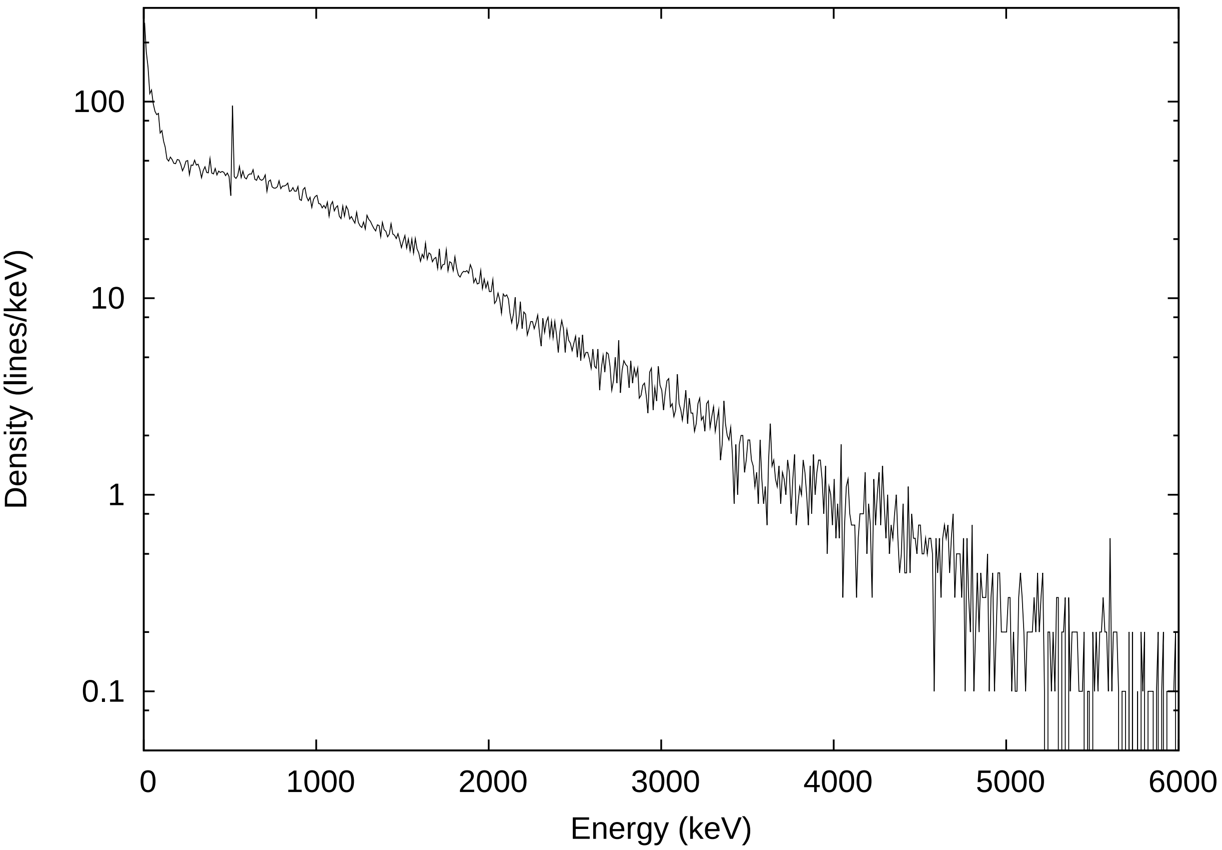

Figure 5.3: Density of gamma-ray lines in the reference library 3.1.1 of Shaman as a function of gamma-ray energy. A constant energy window of 10 keV has been used in the calculation.

and its gamma-ray lines can be read starting from the stated address. When a gamma-ray energy is requested (or more typically a range of energies), the matching energies are searched from the energy index, where the address of the corresponding nuclide in the nuclide index file is stated. The nuclide index is read from the index file, whereafter the nuclide with its gamma-ray lines can be read from the library file.

The table of contents file is in ASCII text format. It includes a list of the nuclides in the library with the number of gamma-ray lines and the header offset in bytes in the library file. In the beginning of this file the total number of nuclides and gamma-rays in the library are stated. There are 2 616 different radionuclides (1 790 gamma-ray emitters) and 81 642 gamma-ray and X-ray lines in the library 2.0.0. The corresponding figures for the library 3.1.1 are 3 648 nuclides (238 stable nuclides, 1 992 gamma-ray emitters) and 80 062 gamma-ray and X-ray lines.

The number of gamma-ray lines per unit energy in the libraries decreases roughly exponentially as a function of gamma-ray energy as can be seen from Fig. 5.3. The significant exceptions are the region below 100 keV where the density is clearly larger due to X-ray lines, and the singular point at 511 keV due to hundreds of annihilation lines in the library. The curve ends at 6 MeV, because there are only 110 gamma-ray lines above this energy in the library 3.1.1, the maximum gamma-ray energy being 11258.9 ± 1.9 keV of ^{20}Na.

```
LU-176    4E+10     Y       B-
57.03     0.95      4         4         71         105
1  0      88.36     14.5      5.86      1.21       0
2  1      201.83    78        0.282     0.164      0
3  2      306.84    93.6      0.0746    0.052      0.9961
4  3      401.1     0.374     0.0347    0.0258     0.0039
W-187     23.9      H       B-
62.0      0.945     10        6         74         113
4  3      7.1       3.4E-3    599.      0          0.033
2  1      72.0      11.9      1.096     1.011      0
1  0      134.3     9.4       2.227     1.72       0.007
2  0      206.24    0.143     3.44      2.60       0
5  2      479.6     23.4      2.05E-2   1.62E-2    0.549
5  1      551.5     5.5       5.66E-3   5.66E-3    0.549
3  0      618.3     6.7       3.03E-2   3.03E-2    0.047
4  0      625.5     1.2       1.04E-2   1.04E-2    0.033
5  0      685.7     29.2      3.23E-3   3.23E-3    0.549
6  0      772.9     4.4       1.83E-2   1.83E-2    0.0423
HG-203    46.6      D       B-
74.6      0.967     1         1         80         123
1  0      279.2     85.1      0.228     0.165      1
```

Figure 5.4: An extract from the decay scheme data file. The file contains the following data: nuclide name, its half-life and decay mode on line 1; KX-energy, fluorescence yield, number of gamma-rays, levels, protons and neutrons on line 2; initial and final level, transition energy, gamma-ray emission probability, total and K-conversion coefficient and percentage of direct decay to initial level on the following lines.

5.4.3 Decay Scheme Data File

In the calculation of coincidence correction, more details of the decay schemes are required than are available in the reference library (indicated with `+`):

- Nuclide data:
  - `+` decay mode
  - `-` KX-energy
  - `+` fluorescence yield
- Transition data:
  - `+` initial and final energy level
  - `-` transition energy
  - `-` gamma-ray emission probability
  - `+` total and K-conversion coefficient
  - `+` percentage branching of direct decay to initial level

A suitable decay scheme data compilation was obtained from STUK (Radiation and Nuclear Safety Authority of Finland) in 1994 [104]. It included decay scheme data of 90 nuclides of interest in environmental monitoring and nuclear power plant supervision, i.e., relatively long lived neutron activation and fission products, uranium and thorium daughters, a few actinides, and common calibration nuclides. We checked the scheme file for inconsistencies and errors with the help of ENSDF-based data compilations [112, 115, 132]. Some new gamma-ray transitions for the existing nuclides and many new radionuclides have been added to the file by us so that it currently includes decay schemes of about 130 nuclides. The format of the decay scheme data file can be seen in Fig. 5.4.

These decay scheme data can in principle be extracted from the ENSDF data base. A program for this purpose has been written [116], but it has appeared that the automatically generated decay schemes cannot be directly used in the coincidence correction calculation. This is because the ENSDF data contain inconsistencies that need to be manually corrected. For example, nuclear states can be grossly imbalanced or the conversion coefficients illogical. The auxiliary program serves to eliminate the most tedious part in creating a decay scheme data file, but its output still needs to be carefully scrutinized by a human expert. Since the program extracts all transitions from the ENSDF, even the weakest ones that have traditionally been omitted when creating the decay schemes manually, the decay scheme file may become unpractically large if all decay schemes available are included in it. Therefore, it may be practical to select only the most interesting nuclides to be included in the file also in the future.

In a standard SHAMAN installation, the decay scheme file resides in the same directory as the reference library files do. Since the file is in text format, it is possible to make corrections and modifications to this file with any text editor, but this cannot be recommended unless the SHAMAN user knows exactly what he/she is doing.

Chapter 6

Case Studies with Expert System Shaman

Besides developing the analysis methods of Shaman and implementing them in practice, the author's focus has been on evaluating its performance. Shaman can be utilized in any application of gamma-ray spectrometry where a complete identification of the spectrum peaks is aimed at. The emphasis of this work has been on airborne radioactivity monitoring, but another recent application of Shaman has been the analysis of pion and proton irradiated samples of the calorimeter and construction materials of the CERN LHC accelerator [XV, 6]. In the latter application, gamma-ray spectra with up to 800 peaks have been encountered, so ordinary air filter spectra with 40–50 peaks can be considered relatively simple. Still, their complete identification is by no means trivial.

The following sections present evaluation results of Shaman in four different cases:

- ordinary air filter spectra measured with a global network of monitoring stations at a prototype stage in 1996–1997,
- air filter spectra measured by the Finnish monitoring network after a minor release from the Novaya Zemlya test site in 1987,
- air filter spectra measured by the Finnish monitoring network after a minor release from the Sosnovyy Bor nuclear power plant in 1992, and
- ordinary air filter spectra measured with a global network of monitoring stations supporting verification of the Comprehensive Nuclear-Test-Ban Treaty (CTBT) in 2007.

Different versions of analysis software and different operation procedures have been applied in these cases, which should be borne in mind when drawing conclusions.

When inspecting identification results for measured air filter spectra as opposed to synthesized ones, the correct results are not known. Nevertheless, a set of nuclides observed in normal conditions is presented in Ch. 3 and the common assumption is that the peaks explained by these nuclides do not need further explanation by other radionuclides. It is not possible, however, to give a probability that a given nuclide is present in a given sample, since it is basically question of an on/off-decision. It is actually quite difficult for a human expert to be consistent and objective in this kind of decisions, giving impetus for developing an automated expert system.

When using a comprehensive reference library, there are several nuclides which could be present in the samples on the basis of the peak data. In the investigations presented in this work, nuclides were neither accepted as correct identifications nor expected to be identified unless there was an acceptable explanation for their presence in air filter samples.

Still, the question remains: when is a peak sufficiently well explained? In Shaman's identification results, this decision is assisted by the *peak explanation share* that is usually presented as a percentage. It tells the share of peak area that is explained by the radionuclides associated with the peak at their calculated activity level. The common range of peak explanation shares interpreted to be a complete identification is 70 %–130 %.[a] The explained share may be out of this range and still accepted as complete for small peaks or peaks fitted inaccurately.

6.1 PIDC Air Monitoring Spectra

The Comprehensive Nuclear-Test-Ban Treaty CTBT was opened for signature in September 1996 and an organization, the Prepatory Commission for the CTBTO, was established for monitoring compliance with the treaty in 1997 (see Ch. 2). Methods for data processing and analysis at the CTBTO were developed at the Center for Monitoring Research (CMR) in Arlington, Virginia, financed mainly by the US government. The CMR hosted the Prototype International Data Centre (PIDC) from January 1995 to September 2001. The PIDC obtained measurement data from prototype monitoring networks of seismic and radionuclide stations. Part of the stations were operated by the PIDC, but many of them belonged to national monitoring networks that voluntarily forwarded their data to the PIDC.

The first operational link between data from an airborne radioactivity monitoring network and expert system Shaman was made during a research project at the PIDC in 1996–97 [VI, VII, VIII, IX]. At the PIDC, Shaman was connected to the radionuclide processing pipeline as an alternative analysis branch and its performance was evaluated in this environment by the author.

6.1.1 Spectrum Analysis at the PIDC

Shaman starts its nuclide identification task from peak search and area determination results from a separate spectrum analysis program. Prior to 1996, Shaman had been using exclusively the spectrum analysis results from Sampo, a software package developed since the 1960's at Helsinki University of Technology [8–15]. The current versions of Sampo run in a PC under MS-DOS, MS-Windows or Linux (see Sec. 2.2.4) and manage every step from

---

[a] A peak may be overexplained, i.e., have an explanation percentage above 100 %, if it belongs to a multiline nuclide whose activity is calculated from all its peaks in the least-squares sense. Due to statistical fluctuations and calibration inaccuracies, the explanation shares of individual peaks vary around the ideal 100 %. The percentage range mentioned above is a rule of thumb based on a long experience.

spectrum acquisition to identification report output. When SHAMAN is used with SAMPO, it replaces SAMPO's identification and reporting parts.

The peak search and fitting results from SAMPO are input into SHAMAN through a dedicated preprocessor program called `idssampo`. At the PIDC, this preprocessor had to be rewritten to read the peak analysis results, produced by the PIDC radionuclide processing software, from an Oracle database. By processing the data from the database into the same format as the data from SAMPO, any changes in SHAMAN's data input could be avoided. This illustrates the portability of SHAMAN: it can take its input from virtually any peak analysis software that is sufficiently sophisticated.

The preprocessor for the PIDC database, known as `db2sha`, was accomplished relatively easily, because the C-functions for exporting data from the database had already been written by the PIDC radionuclide group. It was only a question of applying these functions, making the necessary conversions to the spectrum data, and writing them to a SHAMAN input file in the required format. A few problems with some older PIDC-data were encountered during the development of the preprocessor, and they were solved by implementing some simple checks which prevent the input of incomplete data to SHAMAN. These kind of problems were to be expected in an evolving prototype environment like the PIDC where new concepts were evaluated and constantly developed.

A postprocessor for inputting SHAMAN's results into the PIDC database was also implemented. The postprocessor called `sha2db` reads the identification results from a textual report file produced by SHAMAN, processes the data into data structures used by the PIDC radionuclide library functions, and then calls these functions to write the identification results into the database. The format of the intermediate file was designed to be easily readable with standard C-functions, and still, easily comprehensible for a human user. The latter criterion made debugging and testing of the postprocessor more convenient. The design of this report file was enabled by SHAMAN's report generating toolkit RGL. With RGL, it is possible to tailor identification reports from a brief summary to a most detailed report.

The automated gamma spectrum analysis at the PIDC was accomplished with software that uses algorithms from the commercial Canberra Genie-PC package [135]. Its performance had been tuned into the PIDC environment by the PIDC radionuclide staff, who had tailored the processing parameters and the reference library, in addition to implementing a connection for input data and results into an Oracle database. The performance of this software is discussed in PIDC quarterly reports like Refs. [136] – [139].

It appeared that the PIDC-version of Genie utilized very similar algorithms to those used by SAMPO. The peak search method in both software packages is Mariscotti's generalized second differences method [85]. This was confirmed by searching peaks in several PIDC-spectra with SAMPO using the peak search significance threshold 3.0 that the PIDC utilizes: the resulting peak lists were identical. Also, the nuclide identification algorithm that is used in the Canberra spectrum analysis software was originally developed for SAMPO [10]. This fact was stated in Canberra's documentation.

However, there were major differences between SAMPO and the PIDC-version of Genie, mainly related to peak fitting and baseline determination. In SAMPO, a Gaussian peak with exponential tails on both sides is fitted using shape parameters from a shape calibration (so-called fixed-width fitting), together with a linear or parabolic baseline function. The fitting interval typically spans ten or more channels on both sides of the peak, depending on the spectral structures in its neighborhood. In PIDC-Genie, on the other hand, a pure Gaussian peak function was used and its width was a free variable in the fit (so-called free-width fitting). The baseline function was a constant with a step that was determined prior to peak fitting: the levels of the baseline on both sides of the peak were calculated as the average of three spectrum channels just outside of the peak region.

Peak fitting differences were not quantitatively evaluated during the PIDC-project. Hence, the pros and cons of fixed-width and free-width fitting are described only qualitatively in the following.

Fitting peaks with a free width gives better area estimates for singlet peaks. Fitting unresolved multiplets with a free width may lead to bad area estimates, but if the software cannot resolve close multiplets automatically, this choice may still be better than a fixed-width fitting. On the other hand, the problems with fixed-width fitting are the dependence on the shape calibration quality, which can be assured by making an internal shape calibration for each spectrum, and fitting of naturally wider peaks than ordinary photopeaks (X-ray, annihilation, single escape peaks). The latter problem is difficult to correct in a non-interactive mode, because it requires a different shape calibration for all wide peaks and an automatic decision between these peaks and close multiplets. This may be practically impossible.

The worst difficulties with free-width fitting were experienced with some insignificant Type I peaks.[b] It sometimes happened that a narrow, high peak was fitted into a region where no peak actually existed. With a fixed peak width, the peak area would have remained insignificant, but with a free width, a very large area was occasionally obtained. An additional test should have been implemented in the PIDC peak analysis software to reduce the number of these cases. One way to find this kind of bad fit would have been to compare the fitted area to a channel-by-channel summed peak area and discard the peak if the difference was too large. Nevertheless, results from extensive test runs at the PIDC showed that these and other kinds of Type I peaks did not pose any severe problems to SHAMAN.

6.1.2 Selecting the Test Cases and Processing Parameters

In November 1997, the PIDC radionuclide monitoring network consisted of 20 particulate sampling stations and 2 xenon sampling stations. They were using over 20 different detectors in different combinations, ranging from a single station with two detectors to four stations using a single detector. Almost 200 spectra were processed and reviewed at the PIDC monthly, resulting in an inventory of almost 4 000 reviewed spectra by the end of November 1997. Thus, the PIDC gamma-ray spectrum database provided a great variety of cases for various kinds of investigation.

The operational mode of 24 h sampling, 4–24 h decay, and 20–24 h acquisition was chosen for

---

[b] Type I peaks are also known as false positives, i.e., peaks that are not in fact present, although found by the peak search algorithm. At the PIDC, the peak search was tuned to have as few missing peaks (Type II's or false negatives) as practical, which inevitably led to a few Type I peaks. The average was about two Type I peaks per spectrum [136–139].

the CTBT radionuclide stations during the operation of the PIDC. This mode was known as the *short-cycle* mode. However, many of the stations sending data to the PIDC were operated with a wide variety of modes where sampling and decay times ranged from 1 day to 1 month and acquisition times ranged between 1 hour and 1 week. These modes were known collectively as *long-cycle* mode. The PIDC analysis software was adjusted for the best performance with short-cycle spectra and only the short-cycle spectrum results were reported in PIDC quarterly reports. In order to enable comparisons to the PIDC software, the performance of SHAMAN was assessed for all kinds of spectra and separately for short-cycle spectra only.

In the SHAMAN test runs presented here, the test cases were chosen more or less at random with the goal of having at least one spectrum from each station and detector in the prototype network. Another goal was to have a number of groups of spectra so that the variability in performance could be assessed. This resulted in five groups of 50 randomly chosen spectra, whose analysis results were carefully verified manually [VIII].

In detail, the selection of the test cases was done as follows:

Group 1.
The first group consists of the 23 spectra randomly chosen for the first test set in December 1996 [VI]. This set was expanded with 27 randomly chosen spectra in March 1997 for the evaluation results presented in the second intermediate report [VII]. It contains at least one spectrum from each station and detector, whose spectra had been reviewed at the PIDC since the spring of 1996.

Groups 2–4.
These three groups include the 150 most recent reviewed spectra as of May 28 noon, 1997. They were divided into the three groups in a random manner. The contents of these groups reflected the relative shares of different stations in the PIDC database very well.

Group 5.
The fifth group includes the 50 most recent reviewed spectra as of September 5 midnight, 1997. It also reflected the relative shares of different stations in the PIDC database very well.

The selection of spectra especially into the last four groups can be considered very random. The first group was selected with a specific goal, but still, the representative spectra for each station and detector were picked up randomly, i.e., the first choice was selected in all cases.

The results were obtained with the September 1997 development version of SHAMAN (v. 0.4) using its comprehensive gamma-ray library version 2.0.0 with 2 616 nuclides and over 80 000 gamma-ray and X-ray lines. The parameters of SHAMAN were adjusted for the best aggregate performance with PIDC spectra, and they were kept constant for all analyzed spectra, both short-cycle and long-cycle and even gas sample spectra, with the following exceptions:

- The two escape peak parameters that are used for escape peak area prediction are detector-specific (see Sec. 4.5). These parameters can be determined with one test run with an educated guess for the values, after which the corrected values can be used for all spectra recorded with the same detector. In order to obtain more accurate parameter values, several spectra per detector were used for their determination.

- The spectra of gaseous samples measured at the stations SE001 and RU001 [139, App. 7.1] are essentially free of radon daughters, and therefore, all natural decay chain nuclides were declared as background nuclides in these spectra. In particulate spectra, only the nuclides above radon were declared as background nuclides, and the statuses of the ^{220}Rn and ^{222}Rn daughters were chosen by SHAMAN based on the sample decay time.[c]
- A low-energy cutoff was set to 44 keV for particulate samples, because there were no interesting gamma-ray peaks below the 46.5 keV energy of ^{210}Pb, and X-ray peaks tend to cause difficulties in identification. Moreover, the 44 keV limit was used in the PIDC spectrum processing at the time of the test runs — only some older spectra had peaks below this limit in the database. In the spectra of gaseous samples, the cutoff was set to 30 keV, which is slightly below the 30.6 keV X-ray energy of ^{133}Xe, the lowest energy of interest. A high-energy cutoff was set to 620 keV in gas spectra, slightly above the highest energy of interest (608.2 keV of ^{135}Xe).

From the experiences of the PIDC radionuclide staff and after preliminary test runs of SHAMAN at the PIDC, application-specific expertise was gained concerning some peaks and nuclides encountered in these spectra. This expertise was transferred to SHAMAN to help its nuclide identification task in order to enable using the comprehensive nuclide library.

The first application-specific modification was to apply a user-given explanation to 18 gamma-ray energies frequently seen in the air filter spectra. These peaks belong to two main categories: gammas from isomers and excitation states produced by cosmic rays in the detector and its shield (e.g., 197.9 keV, 803.1 keV), and X-ray sum peaks of ^{220}Rn daughters (e.g., 377.2 keV, 2687.3 keV). It would otherwise have been impossible for SHAMAN to explain these peaks correctly, because they are not included in its library. Prediction of X-ray sum peak areas should actually be possible using the coincidence correction module of SHAMAN, but this feature has not been implemented yet. The current list of air filter gamma-rays with a user-given explanation is presented in App. A.

The second modification was to declare the nuclides ^{71m}Ge, ^{73m}Ge, ^{75m}Ge, and ^{207m}Pb as background nuclides. These nuclides are short-lived activation products, which are occasionally seen in low-level measurements, since they are continually produced by cosmic radiation in the detector or in the lead shielding. Without this modification, these nuclides would have been discarded due to their short half-lives.

The third air filter specific modification was to declare 88 of the most frequently encountered interfering nuclides as unacceptable. This helps SHAMAN to reduce the number of spurious identifications, and at the same time, to reduce the number of false discardings. These nuclides are unlikely to be seen in air filter spectra, and they are irrelevant for CTBT monitoring, so this modification is justifiable. The nuclide list has been modified afterwards

---

[c] The background status of a nuclide has two major effects on SHAMAN's inference: a background nuclide cannot be discarded on the basis of too short a half-life, and the efficiency calibration of background nuclides is corrected for matrix effects [5, p. 54–56].

Group	Number of present nuclides per spectrum	Number of missing nuclides per spectrum	Number of spurious nuclides per spectrum	Nuclide-ID percentage	Spurious-ID percentage
1 (50)	10.06	0.060	0.64	99.40	6.02
2 (50)	9.20	0.060	0.48	99.35	4.99
3 (50)	8.92	0.080	0.28	99.10	3.07
4 (50)	9.44	0.120	0.54	98.73	5.48
5 (50)	9.64	0.160	0.48	98.34	4.82
all (250)	9.45	0.096	0.48	98.98	4.92
mean				99.0 ± 0.5	4.9 ± 1.1

Table 6.1: Nuclide identification in 250 PIDC test spectra by Shaman in September 1997.

as it was by no means considered final in September 1997. The current list of nuclides declared as unacceptable in air filter spectra is presented in App. A.

6.1.3 Results

Shaman's identification results of the randomly chosen 250 PIDC-spectra, divided into five groups, were carefully analyzed, and all peculiarities were recorded to a file in a format suitable for applying basic Unix-commands. The condensed statistics of these test runs are shown in Tables 6.1–6.4 [VIII].

In Tables 6.1 and 6.2, the number of present nuclides is the number of nuclides that is required for a complete explanation of the spectrum peaks. It must be noted that this quantity is not always well-defined. For instance, the nuclides ^{226}Ra and ^{235}U have their primary gammas at 186 keV, and if the other energies of the latter are not seen, it is virtually impossible to divide a peak at 186 keV between these two nuclides. In the analysis of Shaman's results, it was considered sufficient to identify one of these nuclides. If both were identified, they were considered as one nuclide. Another typical ambiguous nuclide is ^{228}Th with a primary gamma-ray at 84.3 keV. This energy is close to the Pb X-ray energy 84.8 keV, which is associated with ^{208}Tl. In many cases, it is impossible to tell if ^{228}Th is needed for the explanation of a peak between 84 and 85 keV, and therefore, Shaman's decision was considered correct in all cases, even if ^{228}Th only had a 5 % share of the peak in some case.

The missing nuclides in Tables 6.1 and 6.2 consist of the nuclides that would explain one or several unidentified or misidentified peaks. Usually these nuclides were identified in other PIDC-spectra measured with the same detector or at the same station. If there was no reasonable explanation for a nuclide, or if it was only associated with an insignificant (Type I) peak, the nuclide was classified as spurious.

In Tables 6.3 and 6.4, the number of found peaks is the number of peaks found and fitted by the Canberra-based PIDC analysis software. However, the preprocessor of Shaman discarded all peaks that were too close to the lower energy discriminator (by default, 8 channels or closer), because these peaks are almost never real. These peaks appeared only in the oldest PIDC spectra, prior to setting the PIDC peak search threshold above the discriminator.

Group	Number of present nuclides per spectrum	Number of missing nuclides per spectrum	Number of spurious nuclides per spectrum	Nuclide-ID percentage	Spurious-ID percentage
1 (25)	9.80	0.040	0.48	99.59	4.69
2 (39)	9.26	0.077	0.54	99.17	5.54
3 (33)	9.18	0.030	0.24	99.67	2.58
4 (32)	9.09	0.094	0.38	98.97	4.00
5 (37)	9.89	0.054	0.35	99.45	3.45
all (166)	9.43	0.060	0.40	99.36	4.07
mean				99.4 ± 0.3	4.1 ± 1.1

Table 6.2: Nuclide identification in 166 short-cycle PIDC test spectra by SHAMAN in September 1997.

Due to this exclusion of peaks, the number of peaks in some spectra may have been different in the PIDC database and in SHAMAN results, but the difference was not significant when the whole set was considered.

The column of unexplained peaks includes both the significant peaks that were left unidentified, and the peaks that were misidentified by SHAMAN. If the peak area was distorted in the peak fitting phase, the correct nuclide may have had a peak explanation percentage far from the ideal 100 %, but the peak was considered correctly identified. Occasionally, SHAMAN's activity calculation resulted in negative activity values for some correct nuclides, and this led to negative peak explanation percentages. This is usually a result of an overdetermined peak-nuclide interference group in the activity calculation and is often an indication of another incorrect nuclide in the interference group distorting the overall solution. It can also indicate errors in efficiency calibration. Peaks with negative explanations were also considered correctly identified, because the association was correct even if the explanation percentage was not. These decisions were analogous to those made by the PIDC radionuclide analysts.

The numbers of insignificant peaks in Tables 6.3 and 6.4 are the most uncertain figures presented, since the distinction between a real peak and an insignificant one is not sharp and calls for a subjective judgment. The critical level is objective in principle, but if the peak area estimate is inaccurate, the critical level test may fail. In these analyses, a peak was classified as insignificant with the following reasoning:

1. SHAMAN indicated the peak to be insignificant. This judgment is made by SHAMAN on the basis of the peak significance, which is a quantity relating the fitted peak area to the baseline level. This quantity is closely related to the minimum detectable activity MDA.The critical level for the insignificant judgment was chosen to be 0.7 (internal SHAMAN parameter) on the basis of preliminary test runs.

2. A peak was unidentified or misidentified by SHAMAN, but there was a Type I -judgment in the PIDC database made by a human analyst. These were mainly cases where the peak area had been overestimated by the spectrum analysis software, or where a non-peak structure had been fitted. The significance test of SHAMAN cannot detect these cases, as it is based on the fitted peak area.

Group	Number of found peaks/spectrum	Number of insign. peaks/spectrum	Number of unexpl. peaks/spectrum	Peak-ID percentage
1 (50)	34.90	2.46	0.40	98.77
2 (50)	38.76	1.80	0.54	98.54
3 (50)	35.14	2.10	0.28	99.15
4 (50)	35.96	1.78	0.62	98.19
5 (50)	37.78	2.32	0.38	98.93
all (250)	36.51	2.09	0.44	98.71
mean				98.7 ± 0.4

Table 6.3: Peak identification in 250 PIDC test spectra by SHAMAN in September 1997.

Group	Number of found peaks/spectrum	Number of insign. peaks/spectrum	Number of unexpl. peaks/spectrum	Peak-ID percentage
1 (25)	43.52	2.20	0.12	99.71
2 (39)	43.85	1.92	0.59	98.59
3 (33)	42.06	1.88	0.09	99.77
4 (32)	42.31	1.50	0.44	98.93
5 (37)	43.38	2.24	0.16	99.61
all (166)	43.04	1.95	0.30	99.28
mean				99.3 ± 0.5

Table 6.4: Peak identification in 166 short-cycle PIDC test spectra by SHAMAN in September 1997.

3. Some insignificant peaks were not indicated by SHAMAN, nor were they judged insignificant by the human analyst during the review of PIDC-software results. However, with the help of visual spectrum inspection and SHAMAN's comprehensive output they could be judged insignificant. These were mainly cases where the insignificant peak had been associated with a common air filter or background nuclide by the PIDC analysis software, in which case the analyst was not expected to comment on the significance of the peak.

The total number of insignificant peaks in the 250 test spectra was 503. The number in the first category was 325, in the second 363 (154 of which not in the first), and in the third 24. Additionally, the 20 peaks that were outside of the user-given energy range in SHAMAN were included in the number of insignificant peaks, because they are handled similarly in the identification percentage calculations. This makes a total of 523 peaks in the insignificant category in 250 spectra, or approximately 2 per spectrum on the average, as shown in Table 6.3.

The *nuclide identification percentages* N, *spurious identification percentages* S and *peak identification percentages* P in Tables 6.1–6.4 were calculated for each spectrum group as follows:

$$N = 100\,\% \times \frac{n_p - n_m}{n_p} , \tag{6.1}$$

$$S = 100\,\% \times \frac{n_s}{n_p - n_m + n_s}\ , \tag{6.2}$$

$$P = 100\,\% \times \frac{p_f - p_i - p_u}{p_f - p_i}\ , \tag{6.3}$$

where

n_p = number of present nuclides,
n_m = number of missing nuclides,
n_s = number of spurious nuclides,
p_f = number of found peaks,
p_i = number of insignificant peaks,
p_u = number of unexplained peaks.

The last equation is analogous to the formula used by the PIDC radionuclide staff [139, p. 31]. It is important to make a clear distinction between the terms peak explanation share (or percentage) that is a quantity related with a single peak and peak identification percentage that is a quantity related with a spectrum or a group of spectra.

6.1.4 Discussion

Averages and Percentages

When the experimental mean and standard deviation of the percentages in Tables 6.1 and 6.3 for the five groups were calculated, the following values were obtained for all 250 spectra in the test set:

$$\begin{aligned} N_{97a} &= (99.0 \pm 0.5)\,\%\ , \\ S_{97a} &= (4.9 \pm 1.1)\,\%\ , \\ P_{97a} &= (98.7 \pm 0.4)\,\%\ . \end{aligned}$$

Omitting the long-cycle spectra and spectra of gaseous samples, the following identification percentages could be calculated for the remaining 166 short-cycle spectra in the five groups of Tables 6.2 and 6.4:

$$\begin{aligned} N_{97s} &= (99.4 \pm 0.3)\,\%\ , \\ S_{97s} &= (4.1 \pm 1.1)\,\%\ , \\ P_{97s} &= (99.3 \pm 0.5)\,\%\ . \end{aligned}$$

It is seen that the identification percentages of short-cycle spectra are slightly better than those for all spectra, although the difference cannot be proven statistically significant with this number of spectra.

The peak identification percentage for short-cycle spectra could be compared to the published PIDC-software percentage [136–139]. Taking into account the first eight monthly peak identification percentages of 1997, when the peak identification percentage showed no significant trends, the PIDC value was $P' = (97.2 \pm 0.5)\,\%$ [140]. However, it must be noted that this value "does not account for peaks that were incorrectly associated to a nuclide" [139, p. 31] in contrast to the presented Shaman figure. The actual value of P for PIDC-identification results would therefore be lower.

Another important difference is that the spectrum set used to obtain the PIDC-percentage was much larger, corresponding to the average of 150 short-cycle full spectra processed at the PIDC monthly. Additionally, the figure corresponded to the operational version of the PIDC-software that was somewhat older and less tailored than the latest development version in September 1997. In any case, an indication of a performance difference between Shaman and the PIDC-software can be seen.

It is non-trivial to show that the observed difference in peak identification percentage is statistically significant. However, the direct comparison in Sec. 7.1 gives a proof of a statistically significant peak identification difference, which gives an indirect proof for the statistical significance of the observed difference in the peak identification percentage presented above.

Another figure that can be compared at this stage is the number of spuriously identified nuclides. Table 6.2 shows that this number was 0.4 for Shaman with the short-cycle spectra and slightly larger for all spectra in the test set. The corresponding figure for the PIDC-software was 0.29 [140], when all the short-cycle spectra processed during the first eight months of 1997 were taken into account. This difference is remarkably small, when one considers the fact that Shaman uses a reference library 100 times larger than the PIDC-software does. A smaller number of spurious nuclides is easily achievable with Shaman by using a smaller library in routine analysis. However, usage of the comprehensive library cannot be ruled out either, because it has the advantage that virtually all incoming spectra are treated as routine cases.

Detailed Investigations on Nuclide Identification

Table 6.5 shows the reasons for the discarding of the 24 missing nuclides in this test set of 250 spectra in September 1997. It can be seen that the most common reason for an incorrect discarding was a missing primary peak (13 out of 24, categories A and B). In these cases, either the primary peak had not been detected by the spectrum analysis software, although one or several secondary peaks of the nuclide had been (cat. A), or the primary peak was out of library lookup tolerance (cat. B). All of these nuclides were close to their detection limit or suffered from calibration inaccuracies.

The number of nuclides in category A is difficult to reduce. These nuclides were discarded due to a completely undetected primary peak. Some of these discardings were due to efficiency calibration inaccuracies, which had led to an incorrectly chosen primary peak. However, some of them illustrated the fact that for some nuclides in some spectra, it is not 100 % certain which peak is most readily detected, when the nuclide activity is only slightly above the detection limit. This kind of cases can possibly be taken into account in Shaman, but this requires some non-straightforward modifications in the code.

It would have been possible for Shaman to identify the nuclides in category B, if a larger library lookup tolerance had been used, but the disadvantages (a larger number of acceptable

Category	Number of nuclides	Explanation
A	7	Secondary gamma-ray(s) had been detected but not the primary gamma-ray
B	6	The energy of the primary peak was out of library lookup tolerance
C	5	The threshold line was too significant: the corresponding peak was not detected although the calculated significance is above detection limit
D	3	The threshold line was too significant: the corresponding peak was detected, but out of the library lookup tolerance
E	3	Nuclide was discarded instead of an interfering nuclide after the activity calculation
F	0	Primary peak was incorrectly explained as an X-ray sum peak or a peak produced by cosmic rays
G	0	Other explanation
total	24	(out of a total of 2 363 nuclides)

Table 6.5: Reasons for the incorrect nuclide discardings in the 250 PIDC test spectra in September 1997.

candidates and, therefore, a larger number of spurious nuclides and an increase in runtime) with this approach outweighed the advantages. Improvement of internal energy calibration methods in either the peak analysis software or in SHAMAN would probably help in reducing the nuclides in category B.

The nuclides in category C were analogous to those in category A, but the discarding had been made due to an undetected significant threshold line, i.e., the most significant line of a nuclide that had not been associated with a peak. Large threshold lines are usually related to inaccuracies in efficiency calibration or spectrum baseline. Thus, category C could be reduced by more accurate efficiency calibrations, which would lead to additional accuracy in gamma line significances calculated by SHAMAN, and by making the modifications in Shaman mentioned above in connection with category A.

The nuclides in category D were analogous to those in category B in the same way as A and C are analogous: the threshold line of nuclides in category D was outside of the energy tolerance. Improvements in internal energy calibration methods would be the solution.

The nuclides in category E had been incorrectly discarded due to difficult interferences in activity calculation. All of these 3 cases corresponded to single-line nuclides, whose identification is always more prone to errors than that of multiline nuclides. Another clear manifestation of this fact is that a vast majority of spurious nuclides are always single-line nuclides that either explain spurious peaks with overestimated areas or peak shares left underexplained by multiline nuclides. In general, an incorrect discarding of a nuclide often leads to a spurious identification.

Category	Number of nuclides	Explanation
I	50	The nuclide was associated with one or two Type I peaks, but no real peaks
J	45	The nuclide was a spurious single-line nuclide, which SHAMAN had not been able to discard with its rule set
K	23	The nuclide explained an almost insignificant fraction of a peak (or several peaks), which had not obtained a 100 % explanation from the correct nuclide(s)
L	3	The nuclide was a spurious multipeak nuclide, which SHAMAN had not been able to discard with its rule set
total	121	(in 250 spectra)

Table 6.6: Reasons for the spurious nuclide identifications in the 250 PIDC test spectra in September 1997.

No incorrectly discarded nuclides were put into categories F and G in the PIDC test spectrum set, but these categories are necessary with other test spectrum sets.

The spuriously identified nuclides were categorized into four categories as shown in Table 6.6. It can be seen that over 40 % of all spurious nuclides had only been associated with false peaks or Type I's (category I). These nuclides were relatively easy to detect when reviewing identification results, as Type I peaks are usually distinctive to the human eye. However, the limit between a significant and an insignificant peak is not a sharp one. A simple study was made among the PIDC radionuclide staff and it showed that the judgments varied from one analyst to the other for these borderline cases. The peak significances calculated by SHAMAN could be used as a guideline when making these judgments.

There is a rule in SHAMAN not to get any candidate nuclides for insignificant peaks, which is based on the peak significances and, thus, on their fitted areas. With the chosen decision level, this test detected about 60 % of all Type I peaks. However, some of the Type I peaks had significances above this level, either due to an inaccurate fitting or due to fitting a backscatter peak or another spectral structure. These latter cases cannot be detected by SHAMAN, unless the decision level is set so high that some real peaks are classified as insignificant, and therefore, there will always be some spurious nuclides associated with insignificant peaks.

The nuclides in category J of Table 6.6 were spuriously identified single-line nuclides. This category included both real and effective single-line nuclides.[d] The great majority of nuclides in categories I and K were also single-line nuclides. If a single-line nuclide passes the tests before activity calculation (most importantly, primary peak found, half-life reasonable, genesis modes compatible), there is actually no way to tell if it is really present or not, so the identifications of the nuclides in category J cannot be considered a severe error.

The 23 nuclides in category K were nuclides with an almost insignificant fraction of a peak

---

[d] An effective single-line nuclide has several gamma lines in the library, but the other lines of the nuclide are insignificant at the primary line activity.

or several peaks. They were accepted by SHAMAN, because they made the explanation percentages of peaks closer to 100 %, i.e., a mathematically more correct solution, but the user could easily see that they are in practice unnecessary for a sufficient explanation of the spectrum peaks. Unfortunately, there is no straightforward way to teach this to SHAMAN, but improving the accuracy of efficiency calibrations and coincidence summing corrections would help to reduce category K.

There were only three nuclides in category L, spurious multiline nuclides, which could not be put into category I or K. This shows that the SHAMAN rules applied before and after activity calculation work well for nuclides with more than one associated peak.

There are two general ways to reduce the number of spurious nuclides in all categories: adding them to the user-given list of unacceptable nuclides, or making a smaller sublibrary. However, it must be emphasized that the results of SHAMAN with the comprehensive library and the parameter settings in use were very good in these analyses, so there is not necessarily a real need for these actions. The power of SHAMAN lies in its capability of using a full library, which enables SHAMAN to give a reasonable explanation to many unexpected peaks in addition to those of the usual suspects. During the one-year SHAMAN project at the PIDC, SHAMAN helped to identify at least the following unusual nuclides in some air filter spectra: ^{41}Ar, ^{59}Fe, ^{77}Br, ^{99m}Tc, ^{108m}Ag, ^{115}Cd, ^{152}Eu, ^{204*}Pb, ^{206*}Pb, and ^{207m}Pb. The less powerful identification algorithm of the PIDC software could not identify them, sometimes simply because the nuclide and its gamma-rays were missing from the reference library.

Detailed Investigations on Peak Identification

The 633 peaks left unexplained or incorrectly explained by SHAMAN could be categorized into 10 groups as shown in Table 6.7. These groups are described briefly in the following:

1. Insignificant peaks (a total of 503) made a majority of all unexplained peaks. These peaks were discussed in more detail above. A majority or 325 of these peaks was flagged by SHAMAN as insignificant, 129 of them were left unidentified, and only 49 of the insignificant peaks were explained with a spurious nuclide.

2. There were 20 peaks below the 44 keV threshold energy or out of the 30–620 keV energy range for gas spectra. SHAMAN did not try to find an explanation for them, because the energy and efficiency calibrations were generally unreliable at their energies. These peaks were subtracted from the total number of peaks analogously to insignificant peaks when calculating identification percentages.

3. Sometimes the free-width peak fitting ended up with an abnormally narrow peak, especially if the peak was small. Therefore, the library lookup tolerance of SHAMAN chose the larger one of the fitted width and the resolution calibration width. With this selection, only 3 peaks were left outside of the lookup tolerance of SHAMAN, when the energy calibration was considered good.

4. There was only 1 peak, for which the fitted width was too large and yet, the energy was out of the lookup tolerance. Abnormally wide peaks were usually explained by

Category	Number of peaks	Explanation
1	503	The peak is insignificant (Type I)
2	20	The peak is below the low-energy cutoff (44 keV in particulate or 30 keV in gas spectra), or above the high-energy cutoff (620 keV in gas spectra)
3	3	The free-width peak fitting has ended up with an abnormally narrow peak
4	1	An abnormally wide peak with an erroneous centroid
5	29	The inaccuracy in energy calibration is too large for the lookup tolerance
6	10	The inaccuracy in efficiency calibration is too large, leading to an insignificant explanation by the correct library gamma-ray
7	17	An abnormally large X-ray peak
8	1	An unresolved multiplet, leading to no association
9	7	Unknown emitter
10	42	The peak is OK, but the correct nuclide has not been identified
total	633	(out of a total of 9 127 found peaks)
total	110	(out of a total of 8 604 real peaks in energy range)

Table 6.7: Reasons for the unexplained peaks in the 250 PIDC test spectra in September 1997.

SHAMAN, because the lookup tolerance of these peaks was the fitted peak width multiplied by a constant. In this case, the baseline was very tilted, which made the peak centroid determination and fitting fail.

5. Excessive inaccuracies in energy calibration were interpreted as the reason for 29 unexplained peaks. The problem had been recognized by the PIDC staff and a new energy calibration update method was already in test use during these analyses at the PIDC. It was expected to reduce the energy-related problems substantially, but its effect was not evaluated during this project.

6. Excessive inaccuracies in efficiency calibration left 10 peaks unexplained. These cases occurred only with a couple of detectors, whose efficiency calibration did not cover the energy range sufficiently well, and only with the background nuclides ^{214}Bi and ^{228}Ac, which have gamma lines over a wide energy range.

7. In spectra from some stations, the 72.8 and 74.9 keV X-ray peaks of Pb were abnormally large in comparison to the gamma-ray peaks of ^{208}Tl, which usually explains these peaks

(the latter together with ^{212}Pb and ^{214}Pb). The probable explanation is that these X-rays were produced in the lead shielding of the detector and detected in addition to the X-rays from the sample. There were a total of 17 cases in this category.

8. There was 1 unresolved doublet which was left unexplained, since the fitted energy did not coincide with the component energies within SHAMAN's lookup tolerance. In other cases of unresolved doublets, at least one of the components was within the tolerance.

9. There were 7 peaks without a known explanation at the time of the test runs. These peaks were at 6 different energies: 98.4 (2 times), 182.5, 261.6, 1014.7, 1021.1, and 1327.1 keV. It is not impossible that some of these peaks were actually insignificant peaks or artificial peak-like structures. However, the 98.4 keV peaks were identified as the $K_{\alpha 1}$ X-ray belonging to ^{234m}Pa and the 1021.1 keV peak as a double annihilation peak in the SHAMAN test runs in May 2003.

10. The last category contained 42 peaks, which had sufficiently accurate energies, widths, and count rates, but which were left unexplained because the associated nuclide was incorrectly discarded by SHAMAN.

The energy lookup tolerance of SHAMAN is based on the fitted peak sigma. In these analyses, the peak-specific lookup tolerance was obtained by multiplying peak sigma with 1.2 and this choice of constant seemed to work well. Of the unexplained peaks, only those in categories 3 and 8, a total of 4 unexplained peaks, could be addressed to a wrong lookup tolerance value. On the other hand, the fact that 90 % of Type I peaks were left unexplained, or were explained with a correctly identified nuclide, shows that the lookup tolerance did not lead to too many candidate nuclides for false peaks. With a wider tolerance, more spurious nuclides would have been associated to Type I peaks. The optimum value of the tolerance constant can be investigated in the future, but the above value is expected to be close to the optimum.

6.1.5 Results with a Newer SHAMAN Version

SHAMAN has been under constant development also after the PIDC-project. Many enhancements have been implemented and corrections have been made. However, there seem not to be changes that would significantly affect the identification results in an automated analysis of routine air filter spectra. The only exception is the reference library that was updated in 2002–2003 [7]. The number of nuclides in the new library is 3 648, including all 238 stable isotopes, and the number of gamma and X-ray lines is 80 062. All known deficiencies in the old library were remedied in the new one.

The set of 250 PIDC test spectra was run through the most recent SHAMAN version v. 1.11 in May 2003. This was possible since the SHAMAN input files produced by the preprocessor `db2sha` had been stored and could be converted to the current preprocessor file format by SHAMAN utility programs. In practice, SHAMAN's input data were unaltered in comparison to the September 1997 test runs.

The identification results of SHAMAN were again inspected carefully. Of course, the results from 1997 were utilized in this process, and a few errors in them were observed. Corrections

Group	Number of present nuclides per spectrum	Number of missing nuclides per spectrum	Number of spurious nuclides per spectrum	Nuclide-ID percentage	Spurious-ID percentage
1 (50)	9.96	0.140	0.68	98.60	6.48
2 (50)	9.10	0.080	0.58	99.12	6.04
3 (50)	8.82	0.080	0.60	99.09	6.42
4 (50)	9.38	0.140	0.66	98.51	6.67
5 (50)	9.52	0.240	0.78	97.48	7.75
all (250)	9.36	0.136	0.66	98.55	6.68
mean				98.6 ± 0.7	6.7 ± 0.7

Table 6.8: Nuclide identification in 250 PIDC test spectra by Shaman in May 2003.

Group	Number of present nuclides per spectrum	Number of missing nuclides per spectrum	Number of spurious nuclides per spectrum	Nuclide-ID percentage	Spurious-ID percentage
1 (25)	9.72	0.000	0.56	100.00	5.45
2 (39)	9.18	0.103	0.56	98.88	5.85
3 (33)	9.09	0.061	0.42	99.33	4.49
4 (32)	9.13	0.156	0.47	98.29	4.97
5 (37)	9.78	0.108	0.62	98.90	6.04
all (166)	9.37	0.090	0.53	99.04	5.41
mean				99.1 ± 0.6	5.4 ± 0.6

Table 6.9: Nuclide identification in 166 short-cycle PIDC test spectra by Shaman in May 2003.

are observable in the numbers of present nuclides and insignificant peaks in the condensed results of Tables 6.8–6.11. One would expect them to be equal with the respective figures in Tables 6.1–6.4, but some minor differences can be seen due to the corrections. A revision of the 1997 results was not warranted, however.

When the experimental mean and standard deviation of the percentages in Tables 6.8 and 6.10 for the five groups were calculated, the following values were obtained for all 250 spectra in the test set in the May 2003 test runs:

$$\begin{aligned} N_{03a} &= (98.6 \pm 0.7)\,\% \qquad [1997\text{: } (99.0 \pm 0.5)\,\%],\\ S_{03a} &= (6.7 \pm 0.7)\,\% \qquad [1997\text{: } (4.9 \pm 1.1)\,\%],\\ P_{03a} &= (99.1 \pm 0.4)\,\% \qquad [1997\text{: } (98.7 \pm 0.4)\,\%]. \end{aligned}$$

Omitting the long-cycle spectra and spectra of gaseous samples, the identification percentages could be calculated for the remaining 166 short-cycle spectra in the five groups of Tables 6.9 and 6.11:

$$\begin{aligned} N_{03s} &= (99.1 \pm 0.6)\,\% \qquad [1997\text{: } (99.4 \pm 0.3)\,\%],\\ S_{03s} &= (5.4 \pm 0.6)\,\% \qquad [1997\text{: } (4.1 \pm 1.1)\,\%],\\ P_{03s} &= (99.7 \pm 0.3)\,\% \qquad [1997\text{: } (99.3 \pm 0.5)\,\%]. \end{aligned}$$

Group	Number of found peaks/spectrum	Number of insign. peaks/spectrum	Number of unexpl. peaks/spectrum	Peak-ID percentage
1 (50)	34.90	2.42	0.44	98.65
2 (50)	38.76	1.80	0.20	99.46
3 (50)	35.14	2.10	0.16	99.52
4 (50)	35.96	1.76	0.36	98.95
5 (50)	37.78	2.30	0.40	98.87
all (250)	36.51	2.08	0.31	99.09
mean				99.1 ± 0.4

Table 6.10: Peak identification in 250 PIDC test spectra by Shaman in May 2003.

Group	Number of found peaks/spectrum	Number of insign. peaks/spectrum	Number of unexpl. peaks/spectrum	Peak-ID percentage
1 (25)	43.52	2.20	0.00	100.00
2 (39)	43.85	1.92	0.18	99.57
3 (33)	42.06	1.88	0.06	99.85
4 (32)	42.31	1.44	0.25	99.39
5 (37)	43.38	2.24	0.16	99.61
all (166)	43.04	1.93	0.14	99.66
mean				99.7 ± 0.3

Table 6.11: Peak identification in 166 short-cycle PIDC test spectra by Shaman in May 2003.

It is again seen that the identification percentages of short-cycle spectra are slightly better than those for all spectra, although the difference cannot be proven statistically significant with this number of spectra. Additionally, the differences between the results of 1997 and 2003 are within statistical uncertainties at one standard deviation level. Thus, there are not any quantitative conclusions to be drawn from this comparison. Shaman's performance in the analysis of routine air filter spectra has remained on a high level.

In addition to the calculated identification percentages, the identification statistics of Shaman can be illustrated as shown in Fig. 6.1 where the distributions of the numbers of unidentified and spurious nuclides and unidentified peaks are displayed. It can be seen that all present nuclides were identified in 223 out of 250 spectra (89 %). In 23 cases (9 %), there was one unidentified nuclide, in 2 cases 2 nuclides, in 1 case 3 nuclides, and in 1 case 4 nuclides. In the worst case, the explanation was a bad energy calibration that distorted the energies of ^{40}K, ^{73m}Ge, and ^{210}Pb photopeaks out of the library lookup tolerance. Additionally, ^{207m}Pb was discarded due to problems in activity calculation. In all 4 cases with more than one unidentified nuclide, the missing nuclides were single-line nuclides that are always the most problematic to be identified.

There were no spurious nuclides in 154 out of 250 spectra (62 %). The number of spurious identifications was 1, 2, 3, 4, and 5 in 61, 18, 8, 6, and 2 cases, respectively. In a worst case, there were 10 spurious nuclides — however, 9 of them were associated with two peaks whose fitting had failed in the PIDC processing. This could be seen from their widths: the sigmas

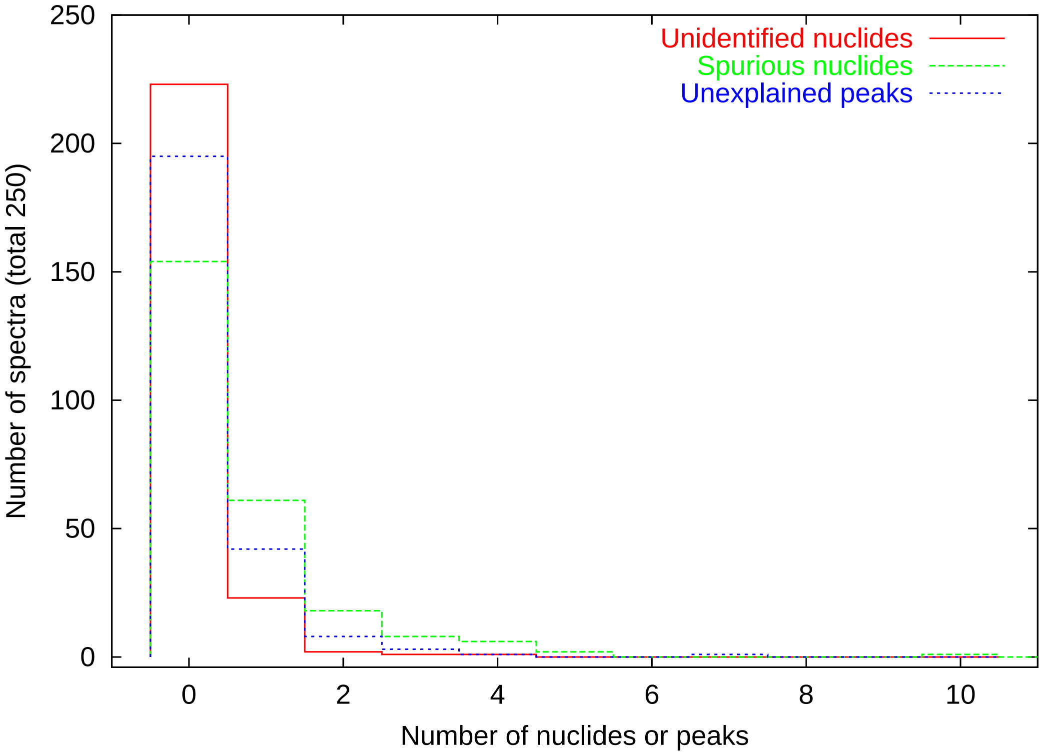

Figure 6.1: SHAMAN's identification statistics for 250 PIDC spectra in May 2003.

of the Gaussian peaks at 277 keV and 301 keV were 3.4 keV and 5.3 keV, respectively, when the sigmas of the neighboring peaks were 1.4 keV. It should be noted that even the correct value of sigma is about twice the accepted value for official IMS stations. Since the library lookup tolerance of SHAMAN is directly proportional to the peak sigma, an anomalously large number of candidate nuclides were accepted for these two peaks and 9 of them happened to pass SHAMAN's discarding rules.

One of the cases with 5 spurious identifications was analogous to the worst case: all 5 spurious nuclides were associated with three peaks with a sigma of 2.9 keV or about double the normal value. The other case with 5 spurious identifications was a gas sample spectrum. The emphasis in PIDC analysis activities was on short-cycle spectra, so the peak analysis or calibrations in this spectrum are likely to be more inaccurate than average, leading to unexplained shares in peaks that SHAMAN attempted to explain with a number of spurious nuclides.

All spectrum peaks that are significant and within the energy range were correctly explained in 195 out of 250 spectra (78 %). There was 1 unexplained peak in 42 spectra (17 %), 2 in 8 spectra, 3 in 3 spectra, 4 in 1 spectrum, and 7 in the worst one. The worst case was the one with a bad energy calibration mentioned above where 4 present nuclides were left unidentified. Also, the spectrum with 4 unexplained peaks as well as two of the spectra with 3 unexplained peaks suffered from energy calibration inaccuracies.

In total, there were 127 spectra out of 250 (51 %) that were identified totally correctly, i.e., with all peaks correctly explained, all present nuclides identified, and with no spurious nuclides. Additionally, there were 67 spectra (27 %) with all peaks correctly explained and present nuclides identified, but with 1–5 spurious nuclides among the correct ones. If we calculate the corresponding figures for short-cycle spectra only, there were 92 spectra out of 166 (55 %) with a totally correct identification and additionally 54 spectra (33 %) with 1–5 spuriously identified nuclides but with all peaks correctly explained and present nuclides identified. This performance level can be considered excellent in automated processing with the full reference library.

6.2 Novaya Zemlya Release in August 1987

6.2.1 Test Material

The number of gamma peaks in the measured spectrum increases drastically in a spectrum measured from a sample containing fresh fission debris. In the test set of PIDC spectra, the samples did not contain almost any anthropogenic nuclides, with the exception of a few calibration and medical isotopes. An average short-cycle particulate spectrum had 43 gamma peaks and 9 nuclides to be identified, as shown in Tables 6.2 and 6.4, but in a fresh fission product spectrum, the number of peaks can be over 100 and the number of nuclides to be identified over 30. This will cause further problems for identification software applying the SAMPO method, especially because nuclide interferences become more abundant, coincidence summing becomes more significant, and weaker gamma lines of many anthropogenic nuclides are detected in addition to their key lines. This will lead to a decrease in the peak identification percentage and an increase in the need for human intervention.

SHAMAN's performance is less affected by the increase of complexity of this kind. This will be shown by processing two sets of spectra containing fission products through SHAMAN: a set of spectra measured by the Finnish monitoring network after a nuclear test in August 1987 in Novaya Zemlya, about 1000 km northeast of Finland, and another set of spectra measured by the same network after a release in March 1992 from the Sosnovyy Bor nuclear power plant, 100 km southeast of Finland. These sets of spectra were kindly made available to us by Dr. H. Toivonen and Mr. A. Leppänen of the Finnish Radiation and Nuclear Safety Authority (STUK) in 1996 and in 1994, respectively.

Gamma-ray spectra of air filter samples collected by the environmental radioactivity surveillance network of STUK are measured in two basic geometries: a Williams geometry and a Marinelli geometry as illustrated in Fig. 6.2. In the Williams geometry, a cylindrical perspex container is placed symmetrically on the top of a detector when the detector is faced upwards. The diameter of the container used by STUK is 42 mm, the height 25 mm and the volume 35 cm^3. The Williams geometry is used in measuring glass fiber filters, which are compressed, ground and pressed into the container before measurement. Sometimes several filters are stacked in a single container and measured simultaneously. [91]

In the Marinelli geometry, a Marinelli beaker is placed on the top of a detector. The Marinelli beaker is a container that has a "well" to fit the detector into, and thus, the detector is efficiently surrounded by the sample. The volume of the Marinelli beaker used by STUK is 500 cm^3 and its approximate dimensions are shown in Fig. 6.2. The Marinelli geometry

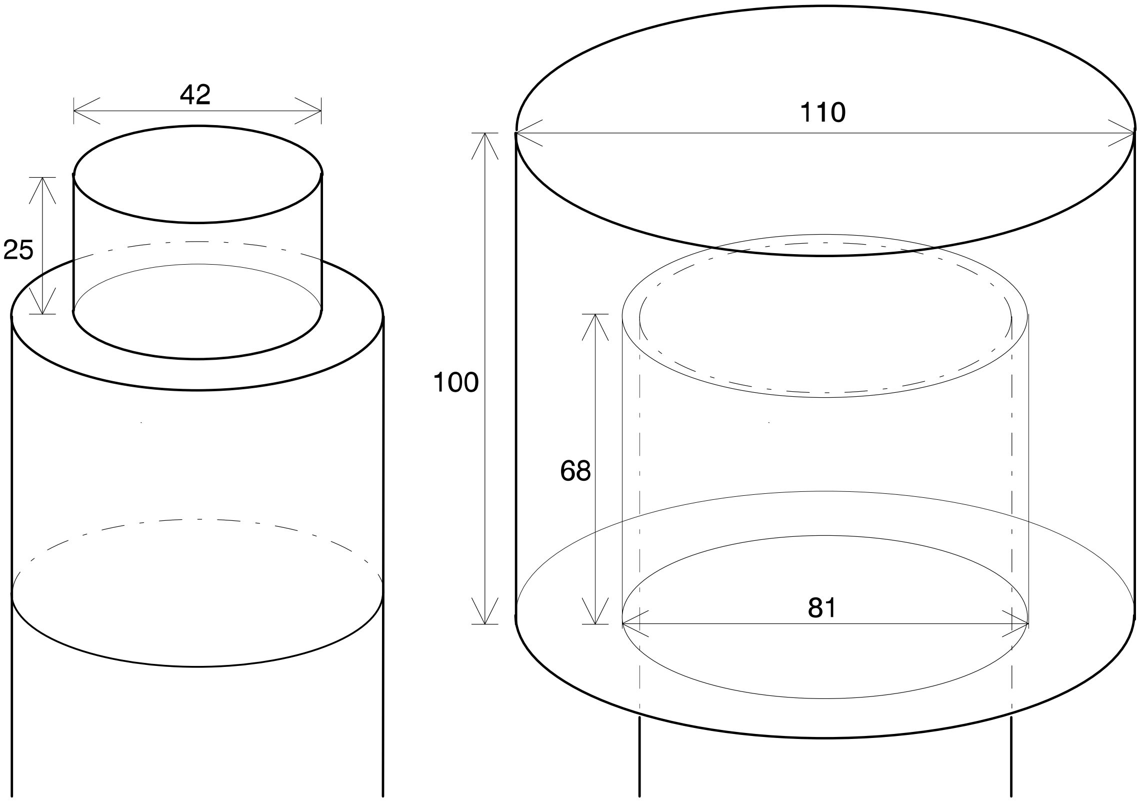

Figure 6.2: The Williams (on the left) and the Marinelli geometry used at STUK in environmental radioactivity surveillance. The approximate dimensions are expressed in millimeters.

is used in measuring filter cartridges of activated carbon that collect gaseous radionuclides in parallel with the particulate collecting glass fiber filters. After sampling, the activated carbon is simply poured into the Marinelli beaker for measurement. [91]

The Novaya Zemlya and Sosnovyy Bor gamma-ray spectra were obtained from STUK in three different file formats, each of which was complemented with a fixed-format text file including calibration file names, time parameters, sample data, and analysis settings used by STUK. Computer programs were implemented to convert these files to the binary spectrum format of SAMPO. Additionally, the energy, peak efficiency and total efficiency calibration files were converted to SAMPO format, and shape calibration files were generated with SAMPO using a representative spectrum for each detector-geometry combination. After the conversion, all spectra were interactively analyzed with SAMPO v. 3.6 whose output was finally converted to the format understood by SHAMAN. The SHAMAN results reported here were obtained in May 2003 using the same version (v. 1.11) as with the PIDC spectra and exactly the same processing parameters, including the comprehensive reference library.

The peak search threshold parameter used in the peak analysis phase in SAMPO was 2.8 for Novaya Zemlya spectra and 4.0 for Sosnovyy Bor spectra. In spite of different values than that used in the PIDC spectrum analysis pipeline (3.0), leading to a different number of peaks found in the search phase, a majority of peaks missed were added in the interactive fitting phase. Therefore, the significance of the largest peaks that are missing from the analysis results should not be too different from that in the PIDC spectra, but of course, *results* *presented for the different groups of spectra are not directly comparable.* The most substantial difference is seen in the share of insignificant peaks: only 0.8 % of the peaks in Novaya Zemlya spectra and 0.5 % of the peaks in Sosnovyy Bor spectra were judged insignificant during inspection of SHAMAN's results. This is because a majority of insignificant peaks were removed in the interactive analysis phase with SAMPO. In the PIDC spectrum results, obtained in an automated pipeline process, this share was of the order of 5 %.

The basic information of the Novaya Zemlya spectra is presented in Table B.1 where the middle of sampling time is used as the sorting criterion. A peculiarity in the spectral data is that the sampling time is not available, only the middle of sampling time and the total air volume pulled through the filter. Therefore, there is no other way to calculate radionuclide concentrations than to apply the counting correction C_c to the start of spectrum acquisition and the waiting correction C_w from that time to the middle of sampling time and to divide the result with the total air volume (see Sec. 4.9). If the sampling time were known, the correction factor required to convert the concentration value calculated this way into the correct average activity concentration during sampling, assuming constant flow rate (Eq. 4.43), would be

$$C_x = \frac{0.5\lambda t_s}{\sinh(0.5\lambda t_s)} ,$$

where λ is the decay constant and t_s is the sampling time. It is seen that $0 < C_x \leq 1$, i.e., the approximate calculation method always overestimates the activity concentration. However, when $\lambda t_s < 1$ or $t_s < 1.44 T_{1/2}$, the correction is below 4 %.

Table B.1 shows that the decay time, or half of the sampling time plus the time between end of sampling and start of spectrum acquisition, is typically 1–4 days and the live acquisition time varies between 6–70 minutes. Thus, the operational mode is relatively far from the short-cycle mode (24 h/24 h/24 h) of the CTBT monitoring network. Especially the short acquisition times will lead to fewer peaks found and hence, fewer nuclides to be identified than would be the case if a 24-hour acquisition time was used with the same samples. This is evident from the identification results: the average numbers of peaks and nuclides are not dramatically larger than in routine air filter spectra.

6.2.2 Results

Averages and Percentages

The identification results obtained with SHAMAN v. 1.11 are presented in Table 6.12 in a similar summary format as the PIDC spectrum results were, but this time for each spectrum separately. It can be seen that the average number of present nuclides increased with 5 and the average number of peaks with 15 from those of the 250 PIDC spectra. Calculating the aggregate nuclide identification percentage, spurious identification percentage and peak identification percentage with the formulas 6.1–6.3 resulted in the following figures:

$$\begin{aligned} N_{NZ} &= 98.0\,\% , \\ S_{NZ} &= 6.6\,\% , \\ P_{NZ} &= 97.0\,\% . \end{aligned}$$

Spectrum identifier	Number of present nuclides	Number of missing nuclides	Number of spurious nuclides	Number of found peaks	Number of insign. peaks	Number of unexpl. peaks
59987f5	21	0	3	60	1	3
42387n3	13	0	0	23	0	1
63887f6	23	0	0	68	1	2
64487f6	16	0	0	45	1	3
59587f5	12	0	1	64	0	0
46087n1	12	0	2	38	0	0
59787f5	12	0	0	46	1	0
64187f6	15	0	1	68	1	0
60587f2	15	2	3	39	0	7
60187f5	17	0	1	75	0	0
60787f2	12	0	2	47	0	7
64787f6	10	0	0	22	0	1
60587f5	14	0	1	68	1	1
61187f2	13	2	1	43	0	3
60687f5	16	2	0	53	0	2
61287f2	13	0	1	42	0	1
61887f2	12	0	0	44	0	0
65387f6	16	0	0	49	0	0
65587f6	12	0	3	62	1	1
61187f5	10	0	0	31	0	0
61987f5	15	1	0	75	1	1
62187f5	14	0	1	35	0	2
62387f5	15	1	1	72	1	5
66087f6	10	0	0	46	0	0
66287f6	15	0	0	52	0	1
62887f5	17	0	0	51	2	3
62787f5	16	1	4	76	1	4
66987f6	13	0	1	55	0	0
67087f6	15	0	0	61	0	0
64287f2	14	0	1	42	0	0
63587f5	16	0	0	40	0	1
65087f2	11	0	1	48	1	0
67687f6	20	0	5	64	1	2
69087f6	16	1	1	64	0	2
total	491	10	34	1 768	14	53
average	14.4	0.3	1.0	52.0	0.4	1.6

Table 6.12: Identification results for 34 Novaya Zemlya spectra by SHAMAN in May 2003.

No uncertainty estimates could be calculated since there was only one set of spectra. However, the values N_{NZ} and S_{NZ} are within the 1σ confidence limit of the PIDC spectrum analysis percentages N_{03a} and S_{03a}. The peak identification is lower than the PIDC value $P_{03a} = (99.1 \pm 0.4)\,\%$, but still on an acceptable level.

A closer inspection of the unidentified peaks revealed that 17 out of 53 unidentified peaks were

Category	Number of nuclides	Explanation
A	2+0	Secondary gamma-ray(s) had been detected but not the primary gamma-ray
B	0+2	The energy of the primary peak was out of library lookup tolerance
C	0+0	The threshold line was too significant: the corresponding peak was not detected although the calculated significance is above detection limit
D	0+0	The threshold line was too significant: the corresponding peak was detected, but out of the library lookup tolerance
E	2+2	Nuclide was discarded instead of an interfering nuclide after the activity calculation
F	2+0	Primary peak was incorrectly explained as an X-ray sum peak or a peak produced by cosmic rays
G	0+0	Other explanation
total	6+4	(out of a total of $121 + 370 = 491$ nuclides)

Table 6.13: Reasons for the incorrect nuclide discardings in the 34 Novaya Zemlya spectra in May 2003. All figures are expressed separately for anthropogenic nuclides and ordinary air filter nuclides.

a result of an inaccurate energy calibration and 13 were abnormally large X-ray peaks. In SAMPO v. 3.6, there was not any automated internal energy calibration feature that the PIDC processing software and UNISAMPO have. The usage of an average energy calibration for all spectra measured with the same detector explained the larger share of energy calibration related problems. The abnormally large Pb X-rays, on the other hand, could be explained by SHAMAN if they were declared as background peaks with the average count rate observed in these spectra. This effort was skipped at this stage, however.

The number of spurious nuclides increased from 0.66 per spectrum in routine PIDC-spectra to 1.0, but this increase was roughly proportional to the increase in the number of peaks found. The spurious identification percentage remained almost constant. This is partly because the share of insignificant peaks decreased, thanks to interactive spectrum analysis. Insignificant peaks would attract spurious nuclides to some extent.

Detailed Investigations on Nuclide Identification

The anthropogenic nuclides that could be identified in the 34 Novaya Zemlya spectra are listed in Table B.2. 18 different man-made radionuclides were observed with varying frequencies in these spectra. Naturally, only the nuclides that could be identified on the basis of the peaks found are listed. A more thorough interactive peak analysis might reveal additional nuclides in some spectra.

The reasons for the incorrect nuclide discardings in the 34 Novaya Zemlya spectra are shown in Table 6.13. The number of missing identifications was so small that no strong conclusions can be drawn. Motivations for the incorrect discardings of anthropogenic nuclides were:

1. ^{125}Sb has four prominent gamma-rays with the following energies and emission probabilities: 427.9 keV / 29.8 %, 463.4 keV / 10.56 %, 600.6 keV / 17.77 %, 636.0 keV / 11.29 % [23]. It was discarded in two spectra because its primary line at 427.9 keV was not found, even if other lines were (category A).

   In the spectrum `61987f5`, the following peaks were seen: 463.1 keV (^{228}Ac explained 32 % of the peak), 600.4 keV, and 636.3 keV (^{131}I explained 53 % of the peak). In the spectrum `62787f5`, peaks at 462.8 keV (^{228}Ac explained 87 % of the peak) and 600.4 keV were found. In both cases, ^{125}Sb was close to its detection limit. SHAMAN's selection of the primary line should be adjusted to handle this kind of cases, but it may be difficult to implement. In interactive analysis, however, it is possible for the analyst to iterate between peak analysis and identification, so the explanation for the peaks of ^{125}Sb should be found with a minor effort — most easily by adding the missing peak at 427.9 keV.

2. ^{132}I was discarded in two spectra because only its primary line at 667.7 keV was found in them and it was explained as the X-ray sum peak of ^{208}Tl that coincides with the ^{132}I gamma-ray energy (category F). In the spectra where ^{132}I was correctly identified, some of its other gamma-ray peaks were found. 772.6 keV seemed to be the secondary gamma line of ^{132}I.

   The ultimate solution for these problems would be to implement the calculation of X-ray sum peak areas in SHAMAN. Currently, the best solution is to keep the ^{208}Tl sum peak explanation despite these problems with ^{132}I, because this results in a correct explanation for the 667.7 keV peak in routine air filter spectra.

3. ^{133}Xe was discarded in two spectra because its peak at 81.0 keV was explained by the spuriously identified ^{223}Ra (category E). This was because the latter nuclide was associated with several peaks whereas ^{133}Xe is a single-line nuclide, and SHAMAN's activity calculation favors multiline associations. This is generally correct, but in this case ^{223}Ra is incorrectly associated to natural peaks and the 81.0 keV peak.

   One solution for these problems would be to add ^{223}Ra to the list of unacceptable nuclides. It is a member of the ^{235}U natural decay chain, but usually only the precursor of this chain is observed in gamma-ray spectra, even in long background measurements. In test runs, declaration of ^{223}Ra as unacceptable was sufficient to make SHAMAN accept ^{133}Xe in these two spectra.

The anthropogenic nuclides listed in Table B.2 were considered reliably identified with the exception of the following two cases:

1. ^{122}Sb was identified only in spectrum `59987f5` where it was associated with a small peak at 564.3 keV. This nuclide has a half-life of 2.7 d which could explain its absence from later samples. Its primary energy is close to the 563.2 keV line of ^{134}Cs, but in this spectrum the peak was much closer to the ^{122}Sb energy. Since the energy calibration seemed to be accurate, the presence of ^{122}Sb could not be ruled out. It is also listed as

Spectrum identifier	Activity of ^{224}Ra (Bq)	Activity of ^{228}Ac (Bq)	Activity ratio
60587f2	18.52	7.711	2.402 (±24 %)
62387f5	166.6	179.8	0.927 (±6 %)
62787f5	184.4	179.9	1.025 (±8 %)
63587f5	20.58	11.66	1.765 (±28 %)

Table 6.14: Activity ratio of ^{224}Ra to ^{228}Ac in 4 Novaya Zemlya spectra.

a CTBT relevant (n,γ)-activation product, because antimony is used to increase the strength of lead in nuclear devices [47].

2. ^{95}Nb was identified in spectrum 62387f5 where it shared a small peak at 766.0 keV with the background nuclide ^{234m}Pa. The latter had its primary line at 1001.1 keV, but ^{95}Nb is a single-line nuclide, so it could not be confirmed with other peaks. Since ^{95}Nb is an abundant fission product seen in many other spectra, it could not be ruled out in this case. However, discarding of ^{95}Nb could not be considered a severe mistake either.

Radium Considerations

^{224}Ra, a member of the natural ^{232}Th decay chain, has been defined as a CTBT relevant nuclide on the condition that it is found at an anomalously high ratio to ^{228}Ac. Such a finding would indicate a test with a ^{233}U weapon [47]. This is why ^{224}Ra is included in Table B.2 among the anthropogenic nuclides. The problem with ^{224}Ra is that its primary and practically only gamma-ray has an energy of 240.99 keV, which is relatively close to the 241.99 keV line of ^{214}Pb. However, the distance is large enough to enable doublet resolving with an up-to-date Ge-detector. Furthermore, the ^{222}Rn daughter chain ^{214}Pb–^{214}Bi dies away from an air filter when a 24-hour decay time between end of sampling and start of acquisition is used as the CTBTO has specified, since it decays with the ^{214}Pb half-life of 26.8 minutes. (This is exactly the motivation for the 24-hour decay time.) The remaining background count rate of ^{214}Pb should not pose a problem if significantly high concentrations of ^{224}Ra are observed.

In the Novaya Zemlya spectrum set, ^{224}Ra was identified in the four cases shown in Table 6.14. In three cases, the activity ratio of ^{224}Ra to ^{228}Ac was 1.0 within 1.5 standard deviations, but in spectrum 60587f2, the difference from unity was 2.5 standard deviations. Ref. [47] does not quantitatively specify the activity ratio to be considered anomalous, but the ^{224}Ra identifications in these spectra were most probably within normal limits. In any case, the ratio varies in a large range depending on the irradiation profile of the ^{233}U production [141].

Cesium Considerations

The fact that this Novaya Zemlya release took place in early August 1987, only 15 months after the Chernobyl accident, complicates the interpretation of the identification results.

Namely, the Chernobyl fallout was resuspended in air and readily detectable by the Finnish surveillance network through the cesium isotopes ^{134}Cs and ^{137}Cs whose half-lives are $T_{134} = 2.066\,\mathrm{a}$ and $T_{137} = 30.07\,\mathrm{a}$, respectively. Being the same element and thus having identical chemical behavior, the activity ratio r of these two isotopes should obey the exponential law:

$$r(t) = \frac{\mathcal{A}_{134}}{\mathcal{A}_{137}} = r_0 e^{-\lambda_r t} ,$$

where $\lambda_r = \ln 2/T_{134} - \ln 2/T_{137} = 0.312\,\mathrm{a}^{-1}$.

An estimate for the activity ratio right after the accident r_0 could be obtained from a wipe sample measurement from April 30 to May 2, 1986 at the Laboratory of Advanced Energy Systems of TKK. According to this measurement, the value was $r_0 = 0.62 \pm 0.03$. Thus, the activity ratio in Chernobyl fallout resuspension should have been $r = 0.42 \pm 0.02$ in August 1987. By inspection of the Novaya Zemlya spectra, this was the ratio observed in all spectra within uncertainties. Therefore, no conclusions should be drawn from the observed activity ratio of the Cs isotopes in these Novaya Zemlya spectra, because there was a substantial contribution from the Chernobyl fallout in their calculated activity concentrations.

6.3 Sosnovyy Bor Release in March 1992

6.3.1 Test Material

Another set of samples containing CTBT relevant radionuclides was collected by STUK in March 1992, when a radioactive release from the Sosnovyy Bor nuclear power plant, some 80 km west of St. Petersburg, could be detected at radiation monitoring stations in Southern Finland [142–144]. The plume of radioactive nuclides was released from Unit 3 of the Sosnovyy Bor nuclear power plant, an RBMK-1000 reactor with the pressurized fuel channel structure characteristic of this model. The release was caused by a breakdown of one of the fuel channels. The reactor was automatically shut down, but some volatile radionuclides were released through the cleaning filters and transported to Finland by southeastern winds. The activity concentrations observed in Finland were of the order of $1\,\mathrm{mBq/m^3}$ or about five orders of magnitude lower than after the Chernobyl accident — clearly detectable by gamma spectrometry but far below the levels that should necessitate countermeasures [142].

The set of Sosnovyy Bor spectra consists of 65 gamma-ray spectra measured in a Williams or a Marinelli geometry as illustrated in Fig. 6.2. 11 different detectors were used in the measurements. The conversion of the spectrum and calibration file formats was done similarly to the Novaya Zemlya spectra. SAMPO v. 3.6 was used in peak analysis that was done interactively and SHAMAN v. 1.11 in identification that was done in a pipeline mode, using the comprehensive nuclide library and exactly the same settings as in the analysis of PIDC test spectra and Novaya Zemlya spectra. Detailed analysis results for the same set of spectra were reported in Ref. [IV] where the 1994 development version of SHAMAN (v. 0.3) was used. The report was utilized in inspecting the results obtained with the newer SHAMAN version.

The spectra had been analyzed and identified at STUK with the GAMMA-83 program which corrects for the true coincidence summing and self-absorption effects seen in these spectra due to close measuring geometries and bulky sources. The identification results had been manually inspected before publication. The most detailed identification reports are those in Ref. [145], which were compared to the identification results obtained with SHAMAN. The agreement was generally very good, when coincidence correction and self-absorption correction was applied in SHAMAN v. 1.11. A clear improvement in this respect was observed in comparison to the SHAMAN results presented in Ref. [IV], mainly thanks to self-absorption correction that was not applied in the old SHAMAN analyses.

The basic information on the Sosnovyy Bor spectra is shown in Table B.3. The decay time in the Sosnovyy Bor spectra varied between 1.5 h–160 d, most typically between 3 h–10 d, and the live acquisition time between 0.6 h–3 d, most typically between 10 h–3 d. The variation in the operational mode was again significantly larger than in the CTBT monitoring network. It can also be seen that some samples had been measured more than once. However, the current version of SHAMAN cannot take advantage of this kind of time slice measurements — it processes each spectrum separately.

Some data in Table B.3 look somewhat doubtful. The sample size of spectrum `13392f5` should probably be 940 m^3 like that of spectrum `07192n4`, because they seem to have been measured of the same sample. Also, the sample size 1 m^3 of spectrum `13692f2` is probably erroneous. It cannot be seen from Table B.3, but calibrations for Williams geometry had been specified in the STUK files for the three activated coal spectra from Loviisa, `07892n1`, `04292f1` and `05392f8`. As a general rule, activated coal samples are measured in a Marinelli geometry, so either the activated coal filter at Loviisa was different from those at the other stations or incorrect calibrations were used for these spectra both by STUK and by us. Anyway, no calibration-related problems could be observed in the identification results of SHAMAN for these spectra.

6.3.2 Results

Averages and Percentages

The identification results obtained with SHAMAN v. 1.11 are presented in Table 6.15 in a similar summary format as the PIDC and Novaya Zemlya spectrum results were. The average number of nuclides per spectrum was slightly larger than that in the Novaya Zemlya spectra, although the average number of peaks was slightly lower. However, this set included the most complicated spectra in all three sets of air filter spectra, with over 30 present nuclides and over 80 peaks in 7 spectra. The aggregate nuclide identification percentage, spurious identification percentage and peak identification percentage calculated with the formulas 6.1–6.3 were the following:

$$\begin{aligned} N_{SB} &= 96.5\,\%\,, \\ S_{SB} &= 10.8\,\%\,, \\ P_{SB} &= 96.6\,\%\,. \end{aligned}$$

Spectrum identifier	Number of present nuclides	Number of missing nuclides	Number of spurious nuclides	Number of found peaks	Number of insign. peaks	Number of unexpl. peaks
13692f2	10	1	0	23	1	1
11592f7	16	0	0	44	1	1
07292k5	9	0	0	20	0	0
15092f7	9	0	0	29	0	0
05392f1	10	0	0	28	6	0
09392f8	13	0	0	32	6	4
04892f1	14	2	0	33	6	2
07392f1	15	0	0	36	6	1
06592f8	16	0	0	41	6	0
15092f6	12	0	0	29	0	1
13592f7	10	0	0	23	0	0
13092f5	7	0	2	75	0	1
07792n4	12	0	1	29	2	0
13492f7	9	0	0	27	0	1
07192k5	9	0	2	24	0	0
08192n3	28	1	13	51	0	2
07892n1	11	0	0	22	0	0
06892n4	31	4	14	84	1	5
07092n4	35	4	11	100	0	8
13692f7	34	2	8	88	0	7
04992f8	9	1	1	78	5	1
20192k1	11	0	3	37	7	1
05092f8	10	0	0	22	3	0
04092f1	6	0	2	71	4	2
08292n3	12	0	0	23	0	0
04192f1	14	0	3	37	6	0
13292f5	44	2	14	148	0	7
07492n4	35	0	7	124	0	3
13192f5	5	0	0	75	0	0
16292f6	19	0	1	45	0	3
16692n3	14	0	0	28	0	0
08192n1	22	1	2	52	4	1
07792n3	6	0	0	26	0	0
20292k1	12	0	0	31	7	1
04292f1	12	1	0	22	2	1
07192n4	37	5	12	105	0	10
13392f5	36	1	3	92	1	3
07592n4	13	1	0	35	0	1
07992n3	16	0	4	34	0	0
08092n3	23	2	3	46	0	2
13692f5	27	0	0	65	1	2
13992f2	10	0	3	26	1	0
15192f6	20	0	1	55	0	1
13592f5	27	2	0	61	2	4
15392f6	12	0	0	31	0	0
14092f7	15	0	0	32	0	0
07292n4	10	0	0	32	0	0
07992n1	10	0	1	24	3	1
13892f2	16	2	1	47	1	1
15492f6	25	1	2	57	0	4
05292f8	12	0	0	33	4	0
05392f8	9	0	1	22	2	1
05592f8	17	0	2	40	5	0
13892f7	8	0	0	33	0	0
05992f8	11	0	0	25	4	1
08892n1	10	0	0	22	0	1
08592n4	17	0	0	41	0	0
07792k5	22	3	3	37	0	3
14392f2	23	0	2	46	1	2
17092f6	23	1	2	54	2	3
16592f6	9	0	0	30	0	0
15892f2	12	0	0	26	0	0
05292f1	14	0	0	32	6	0
15192f5	13	0	0	37	1	2
15592f7	18	0	0	47	0	0
total	1 056	37	124	2 924	107	96
average	16.2	0.6	1.9	45.0	1.6	1.5

Table 6.15: Identification results for 65 Sosnovyy Bor spectra by SHAMAN in May 2003.

No uncertainty estimates could be calculated since there was only one set of spectra, but all three percentages are outside of the 1σ confidence limit of the corresponding percentages for the PIDC test spectrum set. However, there was no severe degradation in SHAMAN's performance, although it apparently had more difficulties with this set of spectra. In an average spectrum, SHAMAN explained 41.9 real peaks correctly and only 1.5 real peaks were left unexplained or incorrectly explained. 0.6 present nuclides were left unidentified while 15.6 of them were correctly identified, in addition to 1.9 spurious identifications.

Our conclusion is that the identification performance of SHAMAN is not significantly deteriorated in the analysis of air filter spectra in a release situation. This is a clear demonstration of SHAMAN's robust design. Even if SHAMAN does a good job even with more complicated spectra, some additional guidance is requested from the human analyst, e.g., in discarding spurious nuclides, in order to make the identification results perfect. The graphical user interface of SHAMAN has been designed to help the analyst to help SHAMAN. After one or two interactive iterations, SHAMAN's results typically reach an excellent level.

Comparison to Old SHAMAN Results

One cannot resist a temptation to compare these results to those obtained with a 9 years older version of SHAMAN. Nuclide identification statistics of the old SHAMAN version are presented in App. B of Ref. [IV]. Calculating the percentages for the set of 65 spectra, the following values were obtained:

$$\begin{aligned} N_{SB94} &= 89.8\,\% \, , \\ S_{SB94} &= 36.4\,\% \, . \end{aligned}$$

The peak identification percentage could not be calculated because information concerning peak explanation was missing, but a clear indication of a boost in SHAMAN's identification performance over the years is observed already by comparing these figures to those achieved with the May 2003 version. Both identification of present nuclides and discarding of spurious ones, naturally interrelated with one another, have experienced a significant improvement.

Actually, the performance difference would be slightly larger if the errors in the data of Ref. [IV] were corrected. The following observations were made when comparing the published v. 0.3 identification results with the identification reports output by SHAMAN v. 1.11:

1. ^{58}Co is not present in the Sosnovyy Bor spectra in contrast to what was presented in Ref. [IV]. In a closer inspection, the 811.2 keV peak in spectrum 17092f6 was judged insignificant (Type I). Actually, SHAMAN v. 1.11 associated two CTBT relevant nuclides, ^{58}Co and ^{156}Eu, with this peak, but they were judged spurious, because the peak was not considered real.
2. ^{152}Eu is not present in the Sosnovyy Bor spectra in contrast to what was presented in Ref. [IV]. Both SHAMAN versions associated it with a peak at 121.4 keV in three spectra, but in all three cases, ^{239}Np had a high activity. A closer inspection revealed that the peak should actually be associated with its $K_{\beta 4}$ X-ray line at 120.969 keV. However, being a minor component in a multiplet, SAMPO had determined its centroid inaccurately enough to push it out of the library lookup tolerance.

3. ^{137m}Ba is not to be calculated as a correct identification in SHAMAN's results in contrast to what was presented in Ref. [IV]. ^{137m}Ba is always present when its precursor ^{137}Cs is, but since the 661.7 keV gamma-ray is included in the list of the precursor's gamma-rays in SHAMAN's reference library, the shorter lived daughter should not be reported separately.

4. ^{226}Ra is not to be calculated as a correct or a spurious identification in contrast to what was presented in Ref. [IV], when both ^{226}Ra and ^{235}U, each with a primary gamma-ray at 186 keV, are present in the background. It is virtually impossible to distinguish these nuclides from each other but in long lasting background measurements. Thus, identification of either or both of these nuclides must be considered equivalent when a peak is present at 186 keV.

5. ^{238}Np is to be considered a correct identification when its primary gamma-ray at 984.5 keV is observed [143], even if it is not listed as a CTBT relevant radionuclide [47, 146]. This identification was made correctly by SHAMAN v. 1.11 in four Sosnovyy Bor spectra. Even the secondary gamma line of ^{238}Np at 1028.5 keV was seen in spectrum `13292f5`, which confirmed this identification. In the other three spectra with a peak at 984.5 keV, the secondary gamma line was below detection limit.

6. The short lived isomers ^{73m}Ge, ^{75m}Ge and ^{207m}Pb are constantly produced by cosmic rays in the detector and its shield. The gamma-rays of these radionuclides were also seen in some Sosnovyy Bor spectra, mainly those with the lowest sample activities. They were correctly explained by SHAMAN v. 1.11 in contrast to v. 0.3.

Detailed Investigations on Peak Identification

In the identification results of May 2003, there were 96 peaks left unidentified or incorrectly identified by SHAMAN. Similarly to the Novaya Zemlya spectrum results, 14 of these peaks resulted from an inaccurate energy calibration because there is no internal energy calibration in SAMPO 3.6. Actually, in two spectra (`15592f7` and `16692n3`) the external energy calibration was so inaccurate that 13 out of 47 peaks and 13 out of 28 peaks had been left unexplained in these spectra, respectively. Therefore, it was replaced with an internal energy calibration in SHAMAN in the analysis results on record. The internal energy calibration was made semiautomatically in the graphical user interface of SHAMAN v. 1.11, thus adding one interactive analysis step to the spectrum processing.

Among the unexplained peaks, there were again 12 abnormally large Pb X-rays at 72.8 keV. They could have been explained by SHAMAN if they were declared as background peaks with the average count rate observed in these spectra, but this was not done at this stage. Other groups of unexplained peaks that would be easy to correct were 8 peaks of ^{140}Ba at 537.3 keV, 8 peaks of ^{97}Nb at 657.9 keV, and 13 peaks of ^{132}I at 667.7 keV. Instead of the correct explanations, the first group of peaks was explained as ^{206*}Pb excitation peaks and the two others as X-ray sum peaks of ^{208}Tl (see Table A.1). These would be the correct explanations for peaks at these energies in routine air filter spectra, but the explanations should be suppressed in more complicated cases like these. With these corrections, the peak identification percentage in the Sosnovyy Bor spectra would exceed 98 %.

It should be noted that the number of insignificant peaks in these spectra was only 15, because the value 4.0 was used as the peak search threshold parameter and some insignificant peaks were discarded in the interactive fitting phase. The majority of peaks in the insignificant peak column in Table 6.15, a total of 92 peaks out of 107, were actually peaks below 38 keV. This energy was given to Shaman as the lowest energy of interest, leading to Shaman omitting any peaks below this threshold. This is the current default setting in Shaman for air filter spectra. In PIDC spectrum analysis, the threshold was at 44 keV, but there is a ^{212}Bi peak at 39.9 keV that should not be excluded from the analysis if observed. Below this energy, there should not be any interesting peaks in air filter spectra. In gas samples, on the other hand, the KX-ray lines associated with Xe-isotopes in the range 29–36 keV should be included in the analysis.

Detailed Investigations on Nuclide Identification

The number of spurious identifications per spectrum was 1.9, almost twice the corresponding value for Novaya Zemlya spectra. However, Table 6.15 reveals that 69 out of the total of 124 spurious nuclides were identified in the 7 most complicated spectra that have over 80 peaks and over 30 present nuclides. Hence, there is a clear correlation between the complexity of the spectrum and the number of spurious identifications. These 7 spectra also showed a sum of 18 missing nuclide identifications of a total of 37 in all 65 spectra. Nevertheless, the nuclide identification percentage was still 93 %, because there was a total of 252 present nuclides to be identified in these spectra. The conclusion is that Shaman's performance with spectra of this level of complication is still good, but the need for interactive analysis increases with an increasing number of peaks and present nuclides.

The anthropogenic nuclide identifications in the Sosnovyy Bor spectra are shown in Table B.4. The maximum number of anthropogenic nuclides is 32 in spectrum `13292f5`, whereas there are 12 spectra without any man-made contribution. The average number of anthropogenic nuclides is 6, but the median value 2 reveals the skewness of the distribution. 34 different anthropogenic nuclides were identified in these spectra with a varying frequency.

The reasons for the incorrect nuclide discardings in the 65 Sosnovyy Bor spectra are presented in Table 6.16, separately for anthropogenic nuclides and ordinary air filter nuclides. It can be seen that Shaman v. 1.11 identified 361 of the total of 393 anthropogenic nuclides. The explanations for the 32 missing identifications were:

1. Anthropogenic nuclides in category A included ^{125}Sb twice and ^{237}U once. The primary and secondary lines of these nuclides typically have almost equal significances. In the problematic cases, they were close to the detection limit and the selection of the primary line happened to fail. Similar cases were seen and discussed in the PIDC and the Novaya Zemlya spectra.

2. The anthropogenic nuclide in category B is ^{105}Rh in spectrum `13692f7`. Its peaks at 306.1 and 318.9 keV are seen in the spectrum, but just outside of the library lookup tolerance. An internal energy calibration would probably help in the identification, since the external energy calibration was off by approximately 0.2 keV in the 300 keV region. When fitting uncertainties were added to this systematic error, these small

Category	Number of nuclides	Explanation
A	3+2	Secondary gamma-ray(s) had been detected but not the primary gamma-ray
B	1+1	The energy of the primary peak was out of library lookup tolerance
C	0+0	The threshold line was too significant: the corresponding peak was not detected although the calculated significance is above detection limit
D	0+0	The threshold line was too significant: the corresponding peak was detected, but out of the library lookup tolerance
E	13+2	Nuclide was discarded instead of an interfering nuclide after the activity calculation
F	14+0	Primary peak was incorrectly explained as an X-ray sum peak or a peak produced by cosmic rays
G	1+0	Other explanation
total	32+5	(out of a total of $393 + 663 = 1\,056$ nuclides)

Table 6.16: Reasons for the incorrect nuclide discardings in the 65 Sosnovyy Bor spectra in May 2003. All figures are expressed separately for anthropogenic nuclides and ordinary air filter nuclides.

peaks were 0.6 and 0.7 keV from the library energies.

3. The number of anthropogenic nuclides discarded by SHAMAN after activity calculation was 13 (category E), a clearly increased share in comparison to the other spectrum sets. This was to be expected: the more complicated spectra, the more complicated nuclide interferences and the more difficult to make the correct discardings. The nuclides in this category were: ^{97}Zr once, ^{99}Mo four times, ^{132}Te twice, ^{135}Xe twice, ^{136}Cs once, ^{147}Nd once, ^{237}U once, and ^{239}Np once. The most important 9 cases are discussed below, the rest are taken as unavoidable casualties.

   (a) ^{97}Zr was discarded in one case where its daughter ^{97m}Nb was identified instead. The primary gamma line of ^{97}Zr is actually produced in the internal transition of ^{97m}Nb to the ground state, but since the gamma is included in the precursor's list, the daughter should not be identified. In 6 cases of 7, ^{97m}Nb was discarded as an intermediate daughter, but for some reason this rule was not used in one case.

   (b) ^{99}Mo was discarded in 4 cases out of 8 due to difficult interferences in activity calculation. The peaks at 181.0, 739.5 and 777.9 keV were associated with several spurious nuclides that SHAMAN was incapable to discard with the current set of rules. Adding a few of these nuclides to the list of unacceptable nuclides would help SHAMAN to identify the correct nuclide.

   (c) ^{132}Te was discarded in 2 cases out of 18 appearances. In these cases, only its 228.2 keV gamma-ray was seen in the spectrum. ^{239}Np has a gamma line at almost the same energy and since it had other peaks in these two spectra, it was identified correctly. In these cases, 34 % and 61 % of the 228.2 keV peak was explained by ^{239}Np, leaving only a fraction of this peak to ^{132}Te. In these circumstances, accepting or discarding of a single-line nuclide tends to become very sensitive to the contents of the interference group.

(d) ^{135}Xe was discarded in 2 cases out of 5 appearances. In both cases, ^{77}Br was identified instead of ^{135}Xe. Their gamma line energies are equal (249.77 keV), but since ^{77}Br was also associated with other peaks, most importantly the 238.6 keV primary peak of ^{212}Pb that is always present in air filter spectra, it was erroneously identified. Actually, ^{77}Br with its half-life of 2.4 days is a common nuisance in air filter spectra, but it cannot added to the list of unacceptable nuclides, because it is a medical isotope. A few ^{77}Br observations were actually made by the prototype monitoring network for the CTBT.

4. ^{97}Nb was discarded 8 times, ^{132}I 3 times and ^{140}Ba 3 times, because their primary gamma-ray peaks were explained as X-ray sum peaks of ^{208}Tl and as ^{206*}Pb excitation peaks as explained above (category F). These nuclides were correctly identified by SHAMAN only when their other gamma-ray energies were seen in the spectrum. Since ^{97}Nb is practically a single-line nuclide, it was never identified by SHAMAN, while ^{132}I and ^{140}Ba were correctly identified in 15 and 7 cases, respectively.

5. ^{137}Cs was incorrectly discarded in spectrum `13692f2`. This was because its activity would have been too high for the sample size of 1 m^3 (category G). However, this value for the total volume sampled was already pointed out as probably erroneous, since this spectrum corresponds to an activated coal sample collected for 50 days. Spectrum `07292n4` was measured at the same station of an activated coal sample collected for 2 days and it had a total volume of 618 m^3 (see Table B.3). Thus, the sample size for spectrum `13692f2` should most probably be above 10 000 m^3. Indeed, when the sample size was manipulated to be above 10 m^3, SHAMAN identified ^{137}Cs correctly.

   SHAMAN calculates a conservative estimate for the maximum activity of each nuclide by assuming that the entire sample mass consists of this nuclide alone. If the calculated and decay corrected activity of the nuclide is above the maximum, it can be discarded. This rule is efficient in discarding short-lived and extremely long-lived candidate nuclides. In air filter samples, the total air volume pulled through the filter is known, but the sample mass has to be deduced from this information. To our knowledge, there is no theoretical limit for how much radionuclides a unit volume of air can contain. In order to make the discarding rule efficient, a maximum mass concentration (g/m^3) has been calculated from the Sosnovyy Bor release spectra. It appeared that ^{137}Cs is the nuclide that had the highest mass concentrations and they were of the order of 10^{-15} g/m^3. On the basis of this observation, a conservative choice for the maximum mass concentration would be 10^{-12} g/m^3. This was SHAMAN's old default value, but for some reason, it has been changed to 10^{-14} g/m^3. The new value would have worked even in this case if the sample size had been correct, so this observation does not give reason for any consequences.

A majority of the identifications of anthropogenic nuclides in the Sosnovyy Bor spectra are certain, since they were made in more than one spectrum and no surprising nuclides

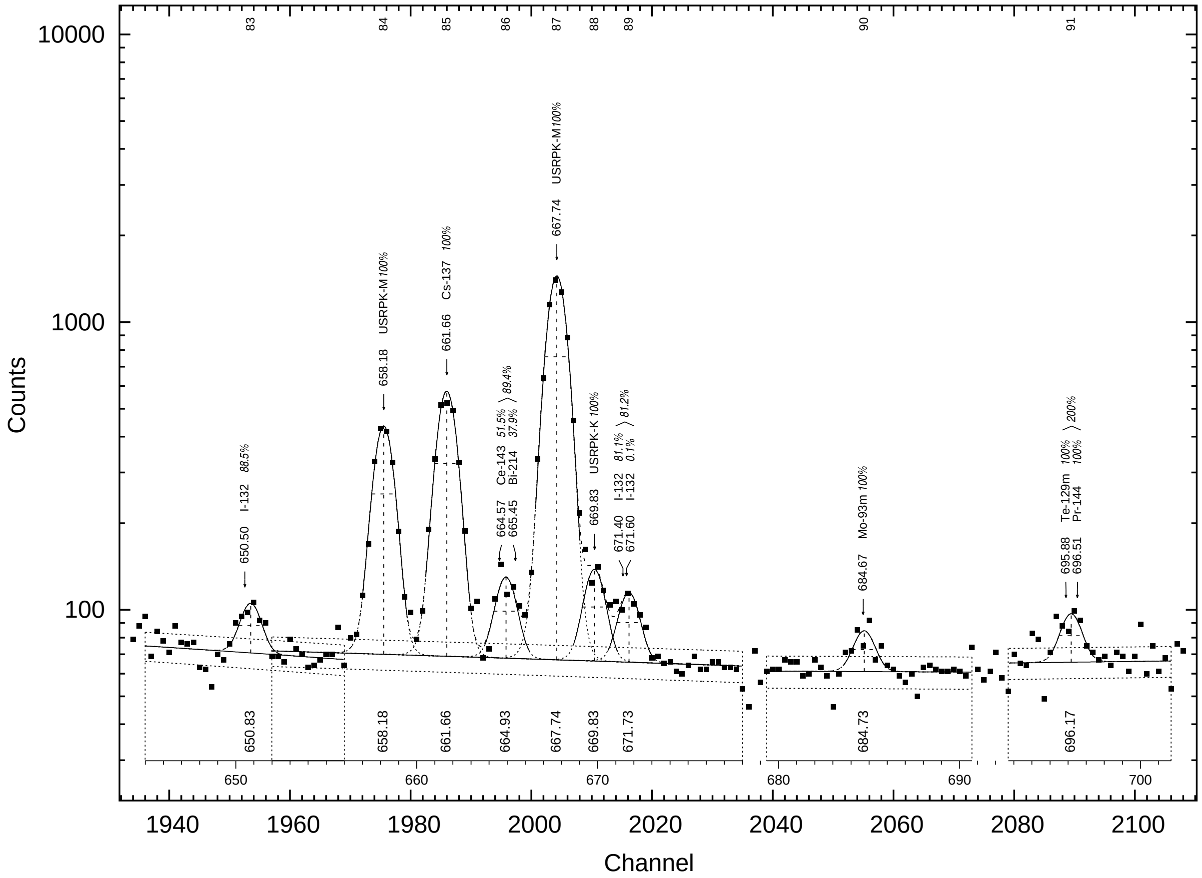

Figure 6.3: Energy region 650–700 keV of spectrum `13292f5` with an interesting peak at 696 keV. SHAMAN's association of the peak with both ^{129m}Te and ^{144}Pr is judged to be correct (see discussion). 8 out of 12 identifications presented in the figure are correct, but the 658.18 keV peak actually belongs to ^{97}Nb (not an X-ray sum peak as indicated by `USRPK-M`), the 667.74 keV and 669.83 keV peaks to ^{132}I (not an X-ray sum peak and a cosmic excitation peak as indicated by `USRPK-*`), and the small peak at 684.73 keV is unidentified.

were identified. The most vague radionuclides are the ones that were identified in only one spectrum, i.e., ^{122}Sb, ^{129m}Te and ^{144}Pr. ^{122}Sb was identified in spectrum `13592f5` where a small peak at 564.0 keV was detected. This nuclide is a CTBT relevant nuclide and since the energy of the peak was only 0.2 keV off of the library energy of the primary line of ^{122}Sb, the identification was considered probable.

The case with ^{129m}Te and ^{144}Pr is intriguing. Both nuclides were identified in spectrum `13292f5` where a peak at 696.17 keV was detected (see Fig. 6.3). The neighboring large peaks revealed that the energy calibration was accurate, so the peak energy was correctly between the 695.88 keV primary line of ^{129m}Te and the 696.51 keV primary line of ^{144}Pr. No other lines of these nuclides were seen to assist in the decision. However, ^{144}Pr ($T_{\mathrm{Pr}} = 17.28\,\mathrm{min}$) could be observed in the spectrum only as the daughter of ^{144}Ce ($T_{\mathrm{Ce}} = 284.91\,\mathrm{d}$) and due to the large half-life difference, the activity ratio of this parent-daughter pair should be 1.000 in equilibrium. If the 696.17 keV peak belonged to ^{144}Pr completely, its activity would be 1.72 Bq (±16 %), but the activity of ^{144}Ce was only 0.683 Bq (±3 %) based on its 133.5 keV

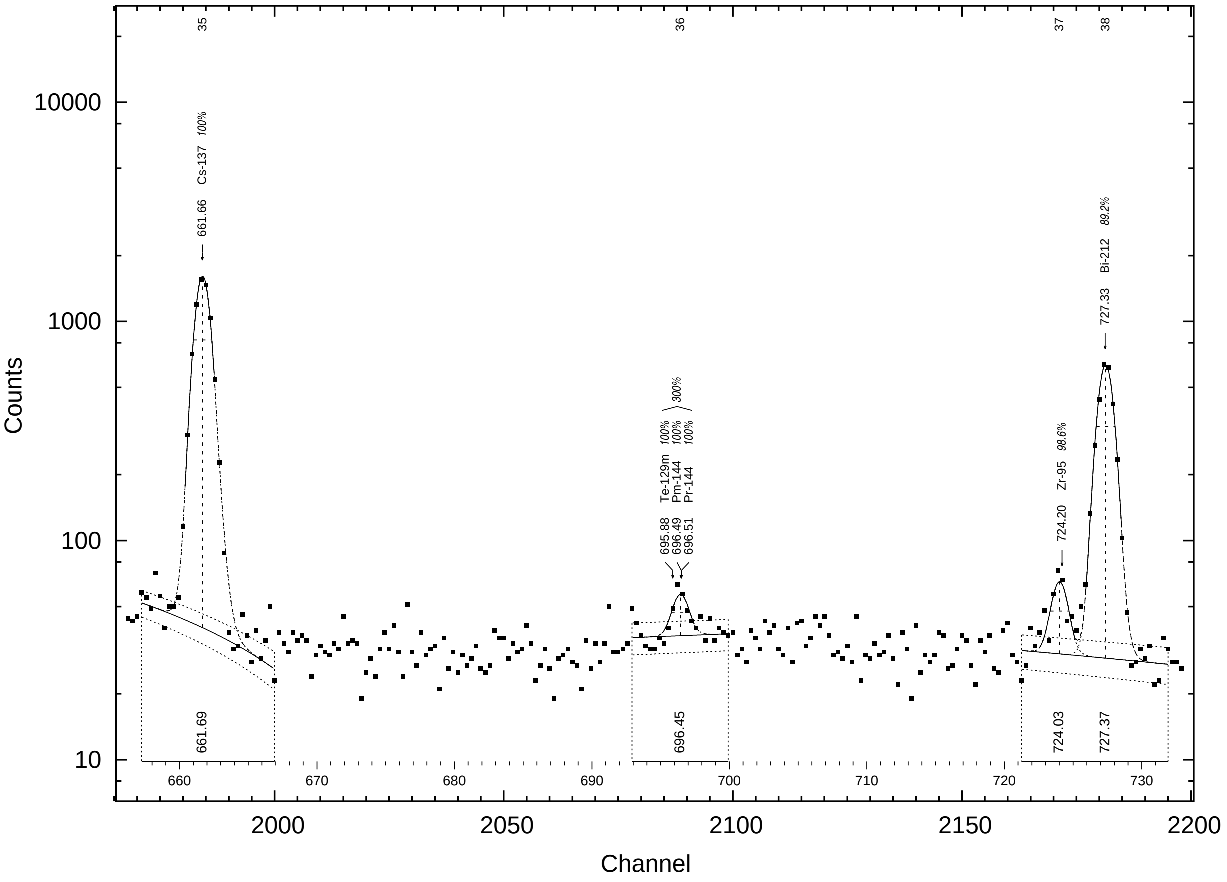

Figure 6.4: Energy region 660–730 keV of spectrum `67687f6` with an interesting peak at 696 keV. SHAMAN's association of the peak with both ^{129m}Te and ^{144}Pr is judged to be incorrect (see discussion). In addition to ^{129m}Te, also ^{144}Pm is spuriously associated with the peak. Other identifications are correct.

peak. Thus, roughly half of the small 696.17 keV peak needed explanation by another nuclide than ^{144}Pr. ^{129m}Te is a CTBT relevant nuclide with a half-life of 33.6 d, so it was a most suitable explanation for the remaining part. In this case, SHAMAN did a perfect job, even if no decay chain or activity ratio related rules were at its disposal. This is because their implementation would be quite involved and application-specific.

It is interesting to note that in the Novaya Zemlya spectrum `67687f6` a peak at 696.45 keV was also explained by SHAMAN with ^{129m}Te and ^{144}Pr. However, the correct identification would have consisted of the latter nuclide only (see Fig. 6.4). This was because the peak energy was within 0.1 keV of the energy of the ^{144}Pr primary line and because the activities of ^{144}Ce and ^{144}Pr were 94 Bq (±4 %) and 114 Bq (±13 %), i.e., sufficiently close to each other to make an explanation with ^{129m}Te superfluous. In this case, SHAMAN also identified ^{144}Pm spuriously. This nuclide decays to the same long-lived nuclide ^{144}Nd as ^{144}Pr does, so it has the exactly same 696.5 keV transition on its gamma-ray list. However, ^{144}Pm is not a fission product, so its probability for being present in this sample is much lower than that for ^{144}Pr.

Spectrum identifier	Activity of ^{134}Cs (Bq)	Activity of ^{137}Cs (Bq)	Activity ratio
11592f7	0.0207 ± 10 %	0.2963 ± 4 %	0.0698 ± 11 %
07392f1	0.0353 ± 12 %	0.4367 ± 5 %	0.0809 ± 13 %
06592f8	0.0303 ± 12 %	0.4017 ± 5 %	0.0754 ± 13 %
08192n3	0.1311 ± 9 %	0.2591 ± 16 %	0.5060 ± 18 %
06892n4	0.1540 ± 9 %	0.2766 ± 12 %	0.5568 ± 15 %
07092n4	0.1669 ± 3 %	0.2937 ± 5 %	0.5683 ± 6 %
13692f7	0.1582 ± 5 %	0.2943 ± 29 %	0.5375 ± 29 %
13292f5	0.2677 ± 3 %	0.4015 ± 6 %	0.6667 ± 7 %
07492n4	0.2686 ± 2 %	0.3885 ± 4 %	0.6914 ± 4 %
16292f6	0.0226 ± 8 %	0.0997 ± 3 %	0.2268 ± 8 %
16692n3	0.0167 ± 35 %	0.2926 ± 3 %	0.0569 ± 35 %
08192n1	0.0362 ± 6 %	0.1629 ± 5 %	0.2223 ± 7 %
07192n4	0.1228 ± 5 %	0.1817 ± 12 %	0.6758 ± 13 %
13392f5	0.1346 ± 4 %	0.2066 ± 4 %	0.6515 ± 5 %
08092n3	0.0671 ± 12 %	0.0967 ± 14 %	0.6937 ± 18 %
13692f5	0.0484 ± 3 %	0.1002 ± 4 %	0.4829 ± 5 %
15192f6	0.0334 ± 11 %	0.0516 ± 7 %	0.6477 ± 13 %
13592f5	0.0353 ± 6 %	0.0645 ± 5 %	0.5470 ± 8 %
08592n4	0.0408 ± 17 %	0.0708 ± 5 %	0.5765 ± 17 %
07792k5	0.0751 ± 8 %	0.5351 ± 5 %	0.1404 ± 10 %
14392f2	0.1260 ± 3 %	0.5901 ± 4 %	0.2135 ± 5 %
17092f6	0.1185 ± 3 %	0.5678 ± 3 %	0.2087 ± 4 %
05292f1	0.0194 ± 26 %	0.2604 ± 5 %	0.0747 ± 27 %
15592f7	0.0193 ± 6 %	0.2405 ± 3 %	0.0804 ± 7 %

Table 6.17: Activity ratio of ^{134}Cs to ^{137}Cs in 24 Sosnovyy Bor spectra.

Radium Considerations

In the Sosnovyy Bor spectrum set, ^{224}Ra was identified four times (see Table B.4). In these cases, the activity ratio of ^{224}Ra to ^{228}Ac varied between 0.9 (±19 %) and 1.5 (±17 %), i.e., it was unity within 2.0 standard deviations. Moreover, ^{224}Ra was identified only in spectra where no anthropogenic nuclides were identifiable, with the exception of ^{134}Cs and ^{137}Cs, so most probably the ^{224}Ra identifications were within normal limits. Of course, this is to be expected since the release came from an ordinary thermal nuclear power plant with ^{235}U fuel.

Cesium Considerations

There is again need for Chernobyl considerations (cf. p. 122). The release from Sosnovyy Bor took place $5\,\frac{11}{12}$ years after the Chernobyl accident, so the activity ratio of the cesium isotopes ^{134}Cs and ^{137}Cs in Chernobyl resuspension should have been $r = 0.62\exp(-0.312 \times 5\,\frac{11}{12}) \approx 0.10 \pm 0.01$ during the Sosnovyy Bor release. Data in Table 6.17, illustrated in Fig. 6.5, show that both ^{134}Cs and ^{137}Cs were identified in 24 of 65 spectra. There were 7 spectra where their activity ratio was close to the Chernobyl value. However, in 13 spectra the ratio was clearly higher: the unweighted average ratio of these 13 cases is $r = 0.60 \pm 0.07$, equal within uncertainties to the Chernobyl value right after release. Apparently, this value is typical of reactor inventory, at least of an RBMK-type reactor.

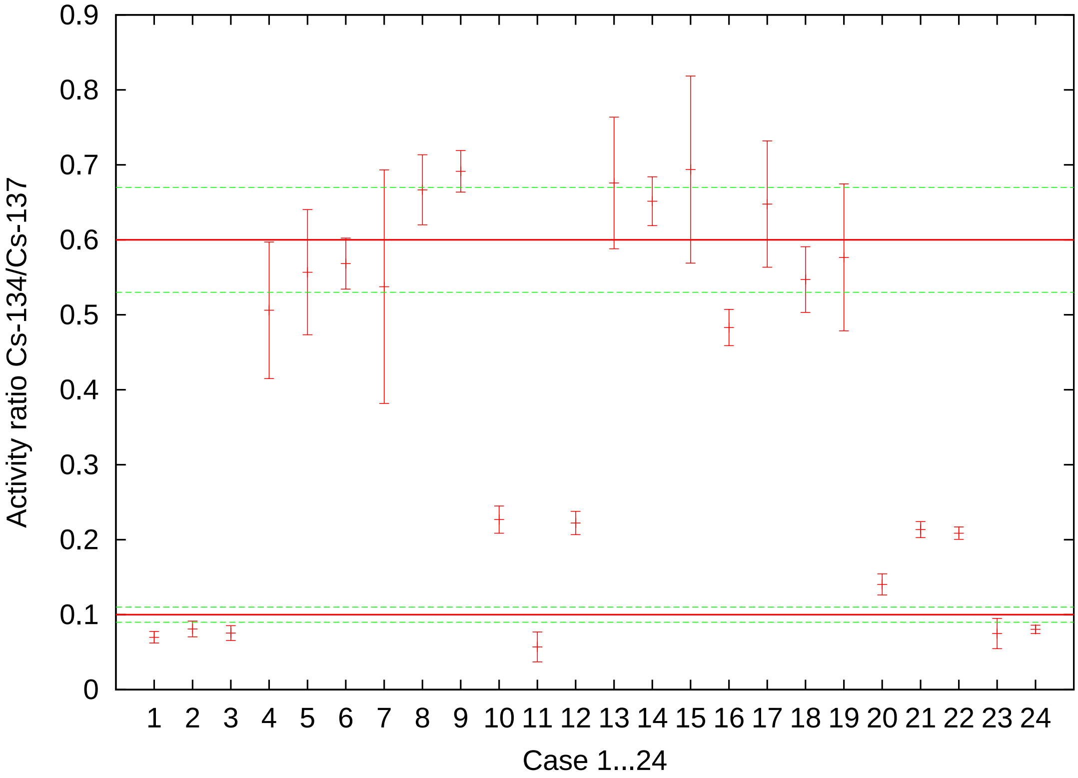

Figure 6.5: Activity ratio of ^{134}Cs to ^{137}Cs in 24 Sosnovyy Bor spectra. The data points are in the same order as in Table 6.17. The ratio for Chernobyl resuspension is 0.10 ± 0.01 and the average of the 13 largest values 0.60 ± 0.07.

There are 4 cases where the ratio is between 0.2 and 0.3. These probably correspond to cases where there is a mixture of Cs isotopes directly from Sosnovyy Bor and from Chernobyl resuspension. This shows that one should be careful in drawing conclusions from activity ratios of different radionuclides. A historical record of observations, in addition to parallel and serial measurements, is essential for correct conclusions.

6.4 IMS Air Monitoring Spectra

In order to quantify the performance of the UniSampo–Shaman analysis pipeline of the Finnish NDC (FiNDC, presented in Sec. 2.2.4), a comprehensive study was carried out in 2000–2001 where the FiNDC pipeline results for IMS spectra were compared to those obtained by the International Data Center (IDC). A complete presentation of the study is available in Ref. [XII] and the most essential findings are presented in Sec. 7.2. At this stage, we only briefly glance at the performance level of the pipeline during the first five months of year 2007.

Table 6.18: Identification statistics by UniSampo–Shaman pipeline for 6 161 full-time air filter spectra from the IMS network in January–May 2007.

Station	Month	Number of spectra	Number of found peaks	Number of insign. peaks	Number of unexpl. peaks	Number of "spurious" nuclides n'_s	Peak id. percentage P' monthly	Peak id. percentage P' for station
RNP01	2007/01	31	1 533	6	65	10	95.7 %	
RNP01	2007/02	27	1 348	8	50	3	96.3 %	
RNP01	2007/03	31	1 352	7	37	6	97.2 %	
RNP01	2007/04	30	1 389	4	50	8	96.4 %	
RNP01	2007/05	31	1 534	11	50	27	96.7 %	96.5 %
RNP03	2007/01	27	1 320	4	55	33	95.8 %	
RNP03	2007/02	27	1 335	9	48	29	96.4 %	
RNP03	2007/03	29	1 464	5	54	36	96.3 %	
RNP03	2007/04	21	1 073	4	43	29	96.0 %	
RNP03	2007/05	21	1 060	2	38	29	96.4 %	96.2 %
RNP04	2007/01	29	1 922	18	81	42	95.7 %	
RNP04	2007/02	28	1 858	18	82	47	95.5 %	
RNP04	2007/03	29	2 049	22	107	47	94.7 %	
RNP04	2007/04	29	2 263	31	164	55	92.7 %	
RNP04	2007/05	31	2 262	15	162	51	92.8 %	94.2 %
RNP06	2007/01	30	1 954	7	72	32	96.3 %	
RNP06	2007/02	26	1 445	6	46	33	96.8 %	
RNP06	2007/03	31	2 155	18	131	90	93.9 %	
RNP06	2007/04	30	2 221	30	175	76	92.0 %	
RNP06	2007/05	30	2 232	50	174	94	92.0 %	94.0 %
RNP08	2007/01	31	519	5	20	14	96.1 %	
RNP08	2007/02	28	381	3	19	8	95.0 %	
RNP08	2007/03	30	501	5	34	9	93.1 %	
RNP08	2007/04	30	364	4	26	11	92.8 %	
RNP08	2007/05	31	436	7	27	7	93.7 %	94.2 %
RNP09	2007/01	31	1 807	12	76	36	95.8 %	
RNP09	2007/02	20	1 078	12	31	40	97.1 %	
RNP09	2007/03	30	1 646	20	70	47	95.7 %	
RNP09	2007/04	30	1 893	30	91	107	95.1 %	
RNP09	2007/05	12	850	13	62	33	92.6 %	95.4 %
RNP10	2007/01	31	2 122	18	97	50	95.4 %	
RNP10	2007/02	28	1 979	18	90	53	95.4 %	
RNP10	2007/03	29	1 998	15	82	43	95.9 %	
RNP10	2007/04	30	2 207	17	108	76	95.1 %	
RNP10	2007/05	21	1 438	12	78	66	94.5 %	95.3 %
RNP11	2007/01	31	1 940	17	86	36	95.5 %	
RNP11	2007/02	28	1 803	10	94	35	94.8 %	
RNP11	2007/03	29	1 927	17	82	43	95.7 %	
RNP11	2007/04	30	2 017	26	106	40	94.7 %	
RNP11	2007/05	31	1 999	12	147	36	92.6 %	94.6 %
RNP15	2007/01	4	48	0	0	2	100.0 %	
RNP15	2007/02	27	773	4	38	13	95.1 %	
RNP15	2007/03	31	1 545	7	51	36	96.7 %	
RNP15	2007/04	30	1 537	8	24	39	98.4 %	
RNP15	2007/05	30	1 366	7	26	28	98.1 %	97.3 %
RNP16	2007/01	31	943	2	50	7	94.7 %	
RNP16	2007/02	28	819	3	34	14	95.8 %	
RNP16	2007/03	31	802	5	42	9	94.7 %	
RNP16	2007/04	30	901	12	62	4	93.0 %	
RNP16	2007/05	30	1 337	10	54	14	95.9 %	94.9 %
RNP17	2007/01	31	963	14	56	17	94.1 %	
RNP17	2007/02	23	678	3	31	9	95.4 %	
RNP17	2007/03	23	674	11	38	4	94.3 %	
RNP17	2007/04	30	949	9	37	7	96.1 %	
RNP17	2007/05	30	1 098	6	40	7	96.3 %	95.3 %
RNP18	2007/01	31	1 693	12	40	33	97.6 %	
RNP18	2007/02	28	1 502	9	26	27	98.3 %	
RNP18	2007/03	17	935	4	18	13	98.1 %	
RNP18	2007/04	15	779	4	13	10	98.3 %	
RNP18	2007/05	30	1 666	14	24	23	98.5 %	98.1 %
RNP19	2007/01	31	1 751	12	50	42	97.1 %	
RNP19	2007/02	28	1 560	10	18	26	98.8 %	
RNP19	2007/03	31	1 750	14	50	30	97.1 %	
RNP19	2007/04	29	1 589	11	24	33	98.5 %	
RNP19	2007/05	31	1 657	4	62	33	96.2 %	97.5 %
RNP23	2007/01	24	1 214	4	25	23	97.9 %	
RNP23	2007/02	14	688	4	40	15	94.2 %	
RNP23	2007/03	30	1 584	11	58	31	96.3 %	
RNP23	2007/04	30	1 696	9	81	38	95.2 %	
RNP23	2007/05	31	1 704	14	90	38	94.7 %	95.7 %
RNP26	2007/01	30	1 187	4	51	15	95.7 %	
RNP26	2007/02	28	990	5	71	15	92.8 %	
RNP26	2007/03	30	939	4	151	19	83.9 %	
RNP26	2007/04	30	1 013	4	158	7	84.3 %	
RNP26	2007/05	31	1 193	6	263	19	77.8 %	86.9 %
RNP27	2007/01	30	1 301	5	56	32	95.7 %	
RNP27	2007/02	28	1 365	11	52	40	96.2 %	
RNP27	2007/03	30	1 510	9	66	65	95.6 %	
RNP27	2007/04	29	1 446	13	63	57	95.6 %	
RNP27	2007/05	30	1 524	4	61	62	96.0 %	95.8 %
RNP28	2007/01	20	642	6	56	8	91.2 %	
RNP28	2007/02	13	103	4	26	8	73.7 %	
RNP28	2007/03	30	215	6	45	8	78.5 %	
RNP28	2007/04	28	724	5	47	20	93.5 %	
RNP28	2007/05	31	1 035	9	47	10	95.4 %	91.8 %
RNP29	2007/01	31	1 025	8	30	19	97.1 %	
RNP29	2007/02	13	457	2	13	12	97.1 %	
RNP29	2007/03	22	745	5	20	18	97.3 %	
RNP29	2007/04	27	705	6	79	21	88.7 %	
RNP29	2007/05	31	1 069	3	34	20	96.8 %	95.6 %

cont. Identification statistics by UniSampo–Shaman pipeline for 6 161 full-time air filter spectra from the IMS network in January–May 2007.

Station	Month	Number of spectra	Number of found peaks	Number of insign. peaks	Number of unexpl. peaks	Number of “spurious” nuclides n'_s	Peak id. percentage P' monthly	Peak id. percentage P' for station
RNP30	2007/01	29	997	11	28	40	97.2 %	
RNP30	2007/02	28	1 050	8	35	36	96.6 %	
RNP30	2007/03	29	1 125	8	44	37	96.1 %	
RNP30	2007/04	30	1 114	7	40	33	96.4 %	
RNP30	2007/05	31	1 150	4	34	38	97.0 %	96.6 %
RNP33	2007/01	24	1 314	7	34	15	97.4 %	
RNP33	2007/02	20	1 165	8	35	18	97.0 %	
RNP33	2007/03	24	1 379	7	41	11	97.0 %	
RNP33	2007/04	18	985	3	32	27	96.7 %	
RNP33	2007/05	24	1 428	6	39	21	97.3 %	97.1 %
RNP34	2007/01	19	338	6	12	11	96.4 %	
RNP34	2007/02	28	605	2	26	5	95.7 %	
RNP34	2007/03	30	602	5	20	13	96.6 %	
RNP34	2007/04	23	338	3	8	5	97.6 %	
RNP34	2007/05	31	730	7	26	9	96.4 %	96.4 %
RNP37	2007/01	31	1 972	6	41	8	97.9 %	
RNP37	2007/02	28	1 855	6	31	4	98.3 %	
RNP37	2007/03	30	1 915	8	40	11	97.9 %	
RNP37	2007/04	30	1 901	7	37	17	98.0 %	
RNP37	2007/05	31	2 010	7	40	16	98.0 %	98.0 %
RNP38	2007/01	29	1 601	18	36	19	97.7 %	
RNP38	2007/02	28	1 565	13	28	13	98.2 %	
RNP38	2007/03	31	1 700	14	26	29	98.5 %	
RNP38	2007/04	30	1 598	26	34	21	97.8 %	
RNP38	2007/05	31	1 699	8	39	23	97.7 %	98.0 %
RNP40	2007/04	26	1 653	12	33	12	98.0 %	
RNP40	2007/05	30	1 881	12	38	8	98.0 %	98.0 %
RNP43	2007/01	21	1 405	13	90	41	93.5 %	
RNP43	2007/02	27	1 761	24	52	42	97.0 %	
RNP43	2007/03	23	1 459	15	40	34	97.2 %	
RNP43	2007/04	30	1 974	33	74	42	96.2 %	
RNP43	2007/05	31	1 997	23	56	52	97.2 %	96.3 %
RNP45	2007/01	31	2 043	12	53	55	97.4 %	
RNP45	2007/02	27	1 840	8	45	51	97.5 %	
RNP45	2007/03	14	926	9	29	20	96.8 %	
RNP45	2007/04	26	1 787	14	41	41	97.7 %	
RNP45	2007/05	28	2 011	17	55	42	97.2 %	97.4 %
RNP46	2007/01	31	1 150	5	15	13	98.7 %	
RNP46	2007/02	27	1 026	5	20	11	98.0 %	
RNP46	2007/03	29	1 110	5	16	8	98.6 %	
RNP46	2007/04	30	1 143	9	18	8	98.4 %	
RNP46	2007/05	31	1 090	8	24	13	97.8 %	98.3 %
RNP47	2007/01	30	1 310	8	28	39	97.8 %	
RNP47	2007/02	26	1 198	5	38	35	96.8 %	
RNP47	2007/03	28	1 195	5	24	20	98.0 %	
RNP47	2007/04	30	1 232	6	31	25	97.5 %	
RNP47	2007/05	31	1 240	7	32	21	97.4 %	97.5 %
RNP49	2007/01	31	1 186	5	30	21	97.5 %	
RNP49	2007/02	26	956	6	32	29	96.6 %	
RNP49	2007/03	30	1 220	6	48	24	96.0 %	
RNP49	2007/04	30	1 255	5	46	42	96.3 %	
RNP49	2007/05	31	1 344	10	41	26	96.9 %	96.7 %
RNP50	2007/01	31	1 280	12	29	23	97.7 %	
RNP50	2007/02	17	750	8	7	7	99.1 %	
RNP50	2007/03	31	1 459	5	28	40	98.1 %	
RNP50	2007/04	30	1 393	7	22	33	98.4 %	
RNP50	2007/05	31	1 339	7	17	17	98.7 %	98.3 %
RNP51	2007/01	30	1 493	11	52	33	96.5 %	
RNP51	2007/02	28	1 437	2	46	30	96.8 %	
RNP51	2007/03	31	1 667	8	65	50	96.1 %	
RNP51	2007/04	30	1 610	16	62	45	96.1 %	
RNP51	2007/05	31	1 842	23	71	58	96.1 %	96.3 %
RNP52	2007/01	31	1 410	10	34	14	97.6 %	
RNP52	2007/02	28	1 354	5	41	15	97.0 %	
RNP52	2007/03	30	1 525	6	36	19	97.6 %	
RNP52	2007/04	30	1 547	8	70	24	95.5 %	
RNP52	2007/05	30	1 717	10	47	24	97.2 %	97.0 %
RNP63	2007/01	30	1 348	7	51	40	96.2 %	
RNP63	2007/02	28	1 218	7	39	39	96.8 %	
RNP63	2007/03	30	1 162	2	43	51	96.3 %	
RNP63	2007/04	30	1 457	6	51	42	96.5 %	
RNP63	2007/05	31	1 529	12	48	58	96.8 %	96.5 %
RNP64	2007/02	1	55	0	0	1	100.0 %	
RNP64	2007/03	31	1 614	11	49	23	96.9 %	
RNP64	2007/04	30	1 453	8	70	13	95.2 %	
RNP64	2007/05	31	1 547	11	48	18	96.9 %	96.4 %
RNP66	2007/01	27	1 445	11	24	18	98.3 %	
RNP66	2007/02	26	1 376	8	17	8	98.8 %	
RNP66	2007/03	31	1 657	9	24	6	98.5 %	
RNP66	2007/04	28	1 478	11	14	12	99.0 %	
RNP66	2007/05	31	1 618	14	28	10	98.3 %	98.6 %
RNP67	2007/03	25	1 152	6	44	21	96.2 %	
RNP67	2007/04	30	1 391	5	83	34	94.0 %	
RNP67	2007/05	30	1 423	6	42	23	97.0 %	95.7 %
RNP68	2007/01	30	984	5	22	7	97.8 %	
RNP68	2007/02	28	943	4	18	5	98.1 %	
RNP68	2007/03	31	1 085	7	31	8	97.1 %	
RNP68	2007/04	30	970	5	20	11	97.9 %	
RNP68	2007/05	31	947	2	25	6	97.4 %	97.6 %

cont. Identification statistics by UniSampo–Shaman pipeline for 6 161 full-time air filter spectra from the IMS network in January–May 2007.

Station	Month	Number of spectra	Number of found peaks	Number of insign. peaks	Number of unexpl. peaks	Number of "spurious" nuclides n'_s	Peak id. percentage P' monthly	Peak id. percentage P' for station
RNP70	2007/01	31	2 225	15	66	44	97.0 %	
RNP70	2007/02	28	1 945	14	48	32	97.5 %	
RNP70	2007/03	31	2 097	12	60	40	97.1 %	
RNP70	2007/04	30	2 039	11	69	50	96.6 %	
RNP70	2007/05	31	2 143	16	64	47	97.0 %	97.0 %
RNP71	2007/01	31	989	5	38	12	96.1 %	
RNP71	2007/02	27	888	7	29	17	96.7 %	
RNP71	2007/03	30	968	6	73	24	92.4 %	
RNP71	2007/04	30	884	5	31	6	96.5 %	
RNP71	2007/05	31	973	5	28	19	97.1 %	95.7 %
RNP72	2007/01	28	1 526	16	36	7	97.6 %	
RNP72	2007/02	28	1 526	9	29	12	98.1 %	
RNP72	2007/03	31	1 623	7	49	11	97.0 %	
RNP72	2007/04	30	1 640	13	38	19	97.7 %	
RNP72	2007/05	31	1 617	8	43	11	97.3 %	97.5 %
RNP73	2007/01	29	1 407	22	43	25	96.9 %	
RNP73	2007/02	28	1 382	13	49	20	96.4 %	
RNP73	2007/03	31	1 521	12	49	20	96.8 %	
RNP73	2007/04	30	1 447	16	54	22	96.2 %	
RNP73	2007/05	31	1 478	15	44	26	97.0 %	96.7 %
RNP74	2007/01	28	1 702	14	64	41	96.2 %	
RNP74	2007/02	28	1 844	7	85	47	95.4 %	
RNP74	2007/03	31	2 005	13	97	67	95.1 %	
RNP74	2007/04	30	1 824	13	60	63	96.7 %	
RNP74	2007/05	28	1 800	15	61	57	96.6 %	96.0 %
RNP75	2007/01	31	1 969	18	58	40	97.0 %	
RNP75	2007/02	28	1 749	16	53	41	96.9 %	
RNP75	2007/03	30	1 906	19	69	47	96.3 %	
RNP75	2007/04	29	1 845	15	64	40	96.5 %	
RNP75	2007/05	31	2 040	20	99	51	95.1 %	96.4 %
RNP76	2007/01	24	1 524	3	41	12	97.3 %	
RNP76	2007/02	28	1 856	13	44	18	97.6 %	
RNP76	2007/03	31	1 969	9	58	16	97.0 %	
RNP76	2007/04	30	1 959	8	35	19	98.2 %	
RNP76	2007/05	29	1 970	3	57	26	97.1 %	97.5 %
RNP79	2007/01	31	2 007	7	33	11	98.4 %	
RNP79	2007/02	27	1 655	10	29	5	98.2 %	
RNP79	2007/03	31	2 005	12	48	11	97.6 %	
RNP79	2007/04	28	1 748	6	48	7	97.2 %	
RNP79	2007/05	30	1 971	11	34	14	98.3 %	97.9 %
average/spectrum			49.66	0.35	1.79	0.96		
total		6 161	305 932	2 150	11 051	5 905		96.4 %

Since the CTBTO utilizes the "peak identification percentage" P', uncorrected for spurious identifications, as a figure of merit, and since this value is easily available in Shaman's identification reports, the performance of the FiNDC pipeline is here quantified in terms of this figure. However, this figure alone is of minor value unless an estimate for the number of spurious identifications is given simultaneously. Strictly speaking, this would require a careful manual analysis or a set of synthesized spectra with exactly known nuclide contents.

For practical purposes, we have developed a simplistic strategy for the assessment of spurious identifications: we assume all spectra to be air filter spectra without anomalous contents, and thus, all anthropogenic nuclide identifications can be considered spurious. Only the following exceptions that are sometimes present in ordinary air filter spectra are always accepted as correctly identified: ^{24}Na, ^{28}Al, ^{60}Co, ^{99m}Tc, ^{131}I, and ^{137}Cs. This method does not give the exact truth, but the spurious identification estimate is realistic and it is easy to calculate in an automated manner.

Uncorrected peak identification percentages P' and average numbers of spurious nuclides n'_s for each operational station in the IMS network during the first 5 months of year 2007 are presented in Table 6.18 where station names have been disguised for confidentiality reasons. There were 42 radionuclide monitoring stations in the IMS network in January 2007. During the spring, 3 more stations came on line.

All full-time spectra (as opposed to preliminary spectra, see p. 7) received by the FiNDC are included in the data set. Ideally, each station but those that came on line during the spring should have sent a total of 151 full-time spectra during the 5-month period. This means $42 \times 151 + 86 + 57 + 93 = 6\,578$ spectra in total. The fact that 6 161 spectra were received shows that the average availability of the IMS stations was about 94 % during this time. Of course it is possible that some spectra were not forwarded from the IDC to the FiNDC and that some full-time spectra were incorrectly named in the FiNDC processing pipeline, but these failures are not very probable. The software versions in the FiNDC pipeline remained unchanged for this time period: UniSampo v.2.44 and Shaman v.1.15.

The figures are best comparable with those in Tables 6.9 and 6.11, obtained for 166 short-cycle PIDC spectra using the peak analysis data from the PIDC software and a slightly older version of Shaman. It can be seen that the average number of found peaks per spectrum is larger by 7 in the FiNDC results. This indicates a substantial difference between the spectrum sets due to different locations of stations and different detectors. However, a significant part of the difference can be addressed to different values of the peak search threshold parameter: 2.4 at the FiNDC vs. 3.0 at the PIDC.

The number of insignificant or Type I peaks is substantially lower in the FiNDC results. Of course we must remember that the PIDC results were manually reviewed and it was found that Shaman could indicate roughly 60% of the insignificant peaks. The number of peaks categorized automatically as insignificant or out of energy range by Shaman was actually 147 in 166 short-cycle PIDC spectra or 0.9 per spectrum. This is still a factor of 3 larger than the FiNDC value obtained in an analogous manner. The lower value is thanks to the peak significance test that is made in UniSampo after peak search.

However, it seems that many of the unexplained peaks in FiNDC results are actually insignificant peaks that escape the significance tests in UniSampo and Shaman, because their areas have been overestimated. The distribution of significances of unexplained peaks, an average of 1.8 per spectrum in the FiNDC results, is very much concentrated on low values, close to the decision limit. Therefore, it is probable that the share of Type I peaks among the unexplained peaks is significant in the FiNDC results. This means that there is still room for optimization of especially the peak analysis parameters of UniSampo, but probably also for tuning of Shaman's parameters.

In any case, the resulting average "peak identification percentage" P' is significantly lower for FiNDC results than the previously observed value for short-cycle PIDC spectra: 96.4 % vs. 99.7 ± 0.3 %. If we calculate the PIDC spectrum percentage similarly to the FiNDC value, i.e., on the basis of the uncorrected Shaman results, we obtain $P' = 98.6$ %. This value is lower than the correct value, because the effect of unexplained Type I peaks, lowering the percentage, outweighs the opposite effect of spuriously identified peaks. The FiNDC value is still significantly lower than this value.

Month-to-month variations in the figures of Table 6.18 are not very large, whereas the variations from one station to another are substantial in some cases. Especially the identification results for stations RNP26 and RNP28 seem to be inferior to those for other stations. The explanation in both cases is hardware problems. The station RNP26 experienced anomalously large gain drifts in March-May 2007. Apparently, the problem was repaired in late May, because the monthly peak identification percentages have been above 97 % since June 2007. The station RNP28 sent several empty or near-empty spectra in February-March 2007, but that problem was also repaired and P' returned to a level comparable with other stations.

These observations show that even if the accuracy of the peak identification percentage can be questioned, it is a useful parameter for quality control. It is well complemented by the estimate for spurious identifications.

The estimated number of spurious identifications in the FiNDC results is $n_s' = 0.96$ per spectrum on the average. A similarly obtained estimate for the short-cycle PIDC spectra is $n_s' = 0.55$. (The correct value is $n_s = 0.53$ as shown in Table 6.9. This proves that the estimation method works quite well.) The substantial increase is addressed to the larger number of peaks found, leading to a larger number of candidates and a larger number of spurious nuclides remaining in the results. The optimization of the processing parameters mentioned above will probably help in reducing the spurious identifications. Additionally, the list of spurious identifications in these IMS spectra could be inspected and the list of unacceptable nuclides in air filter spectra (Table A.2) updated accordingly.

It should be noted that the IDC processing software, a newer version of the PIDC software, also shows a degraded performance with the IMS spectra when compared to its performance with the PIDC spectra. A direct comparison in Sec. 7.2 shows that the performance difference between SHAMAN and the IDC processing software has actually increased since 1997.

Chapter 7

Comparison of SHAMAN and Other Identification Systems

Benchmarking with previously existing methods is frequently required when developing scientific procedures and techniques. If there is no other suitable benchmarks, at least the method should be compared to its previous versions, the baseline. Creating suitable tests is a question of experimental design and quality control, and it is also applicable to development of scientific software.

When comparing a new version of a software package to its previous versions, a standard test set should be created. In SHAMAN's case, a set of 50 measured and synthesized spectra was collected already in 1992. The test set was later abandoned due to two reasons: the majority of the measured spectra in the set were too simple to show a difference and the synthesized spectra in the set had been generated with partially obsolete data of Erdtmann-Soyka [88, 89]. The incorrect peaks in the latter ones led to artificial identification problems for SHAMAN. Instead, a set of 87 air filter spectra and another set of 30 more complicated spectra from different applications have lately been utilized in comparing new versions of SHAMAN to previous ones. Especially the former test set is quite easy for the current version of SHAMAN, so it should probably be enlarged to distinguish significant differences between software versions from inevitable differences due to rounding errors etc.

When comparing a software package to other software packages, one should use a similar approach: a standard test set with sufficient versatility. However, the situation is more complicated than when comparing different versions of the same software package. Special effort should be put into leveling the playground as much as reasonable. This means that all software packages should be tuned to the application and the test material should be compiled in a random manner. Yet, the objectivity of the comparison is always questioned when "own" software package is compared to "someone else's" software package, especially if the former comes out as the winner of the comparison.

The basic prerequisites for this kind of direct comparisons are:

1. The performance of all software packages is tuned to the application, if not strictly optimized.

2. The test material is chosen among relevant cases in a random manner but with a full coverage of the application.

3. All relevant comparisons are made and they are made in an objective manner.

When these assumptions are fulfilled, it is sufficient to compare the final results given by the software packages, disregarding the means that have been used to reach the results. Still, the comparison is valid only for the actual versions of each software package.

Intercomparisons of software for gamma-ray spectrometry have been organized by the IAEA and also by other organizations [147–153]. In these intercomparisons, the emphasis has been on the peak analysis, i.e., peak finding and quantification. Of course, nuclide identification and quantification is based on peak analysis, so it must first be in statistical control before the latter phases can be evaluated. Ref. [151] also contains an evaluation of nuclide identification where an air filter sample containing Chernobyl debris is under scrutiny. However, the intercomparison has been made with tailor-made reference libraries of only 24 radionuclides, so its applicability is quite limited.

In the following, we present two comparisons between identification results of the expert system Shaman and those from another software package based on the Genie software package. The first intercomparison was made at the Prototype International Data Centre (PIDC) in 1997 and it truly measures the difference in nuclide identification, since the peak analysis results are input identically for both identification software packages. The second intercomparison was made at the Finnish National Data Centre (FiNDC) in 2001 and it compares two automated analysis pipelines where the analysis software packages differ also on the peak analysis part.

7.1 Comparison between Shaman and the PIDC Nuclide Identification

The nuclide identification part of the PIDC-software, which is based on Canberra's Genie-2000 Spectroscopy Tools software package and referred to as Genie-NID in the following [135], provided an obvious benchmark to compare Shaman's performance to in 1997. However, a direct comparison between Shaman and Genie-NID was not very straightforward for several reasons. First, the gamma libraries, which are essential for the identification, differed substantially in size and in contents [VI]. Second, the nuclide identification methodology of Shaman [5] is very different from that of Genie-NID, which is based on algorithms originally developed for Sampo [10]. And third, the performance of both software is very much dependent on several parameter settings.

However, the basic prerequisites formulated above were met: the PIDC-software had been tuned to the application by the PIDC radionuclide staff and Shaman by me. The test set was selected randomly and all comparisons were made objectively. The comparisons were divided into qualitative and quantitative ones. In the former, differences in peak and nuclide identification were investigated on the level of peak-nuclide associations. In the latter, differences in calculated nuclide activities and related quantities were assessed.

It should be noted that both software packages were under continuous development before and after the comparison, so the performance difference has probably changed after the comparison. However, the comparison gave a clear indication of the performance difference between these two software packages in the fall 1997. Later comparisons, e.g., the one presented in Sec 7.2, have actually confirmed that the difference has remained close to the observations presented here.

7.1.1 Qualitative Differences in Identification Results

In the fall 1997, Shaman was run at the PIDC in an experimental analysis pipeline where the calibrations and peak analysis were made identically to those in the official development pipeline of the PIDC radionuclide group. The identification results of the development pipeline, obtained with Genie-NID, were also saved in the Shaman-database. Parallel database tables for identification results were filled with Shaman's results. This procedure enabled the comparison of identification results with simple database queries in one database. It also ensured that the results saved in the database had always been obtained with the latest versions of both software packages. The Shaman-database included about 700 spectra as of the end of November, 1997.

Some 5–10 new spectra were sent to the Shaman-database daily by a data subscription mechanism that had been provided by the PIDC radionuclide group. These subscriptions sent full spectra (as opposed to preliminary ones) from the six short-cycle stations in the prototype network (AR001, AU001, CA002, KW001, SE001, and US001), which operated on a daily schedule [139, App. 7.1]. In addition to these spectra that corresponded to the operating schedule of the final network, and thus, were the most relevant ones, the full spectra from two long-cycle stations (CA001 and FI001) and from one gas sampling station (SE001) were automatically sent to the database. The analysts also had the possibility to send any interesting cases from other stations manually to the database.

The most recent spectra from each station with a valid subscription were chosen for the comparison of identification results. However, the gas sample identifications could not be compared, because Genie-NID identified gas spectra with a small library containing only the interesting Xe-isotopes. The number of spectra was chosen to be 50 for each short-cycle station but AR001 and KW001, and 25 for AR001, KW001, and the two long-cycle stations, leading to a total of 300 spectra ($4 \times 50 + 4 \times 25$). This limitation was necessary due to the limited number of spectra in the database, but in any case the number of spectra proved to be sufficient for statistically significant conclusions. The chosen spectra were the most recent ones as of November 10, 1997, so the spectrum set was randomly compiled. However, the following exceptions were necessary:

1. The CA002 station had suffered from a hardware failure on October 28, 1997. The CA002 spectra measured after this date had to be excluded from this comparison due to their severe distortions.
2. Some of the ordinary air filter spectra from KW001 and US001 had been replaced with test spectra containing anthropogenic nuclides during the second half of October and the first half of November in 1997. These spectra were excluded from this comparison.

All spectra in the SHAMAN-database were re-processed with the most recent versions of `rms_analyze` (dated November 25, 1997) and `rms_shaman` (dated November 7, 1997), utilizing the most recent processing parameters with both software packages. The raw data for the comparisons were collected with an SQL-script, which reported the following lists for selected analyses:

1. identically identified peaks,
2. differently identified peaks,
3. peaks identified by SHAMAN alone,
4. peaks identified by Genie-NID alone, and
5. peaks left unidentified by both software.

The following data were manually compiled from the SQL-output into a file:

1. Number of peaks identified by SHAMAN, not by Genie-NID. These peaks were further divided into two categories: significant and insignificant peaks.
2. Number of peaks identified by Genie-NID, not by SHAMAN. Division into significant and insignificant peaks was impossible, since Genie-NID did not make the distinction.
3. Number of spurious nuclides identified by SHAMAN.
4. Number of spurious nuclides identified by Genie-NID.

The manual data inspection was simplified by the following assumptions:

1. All peaks associated with common air filter nuclides were associated correctly by both software.
2. All unfamiliar nuclides were considered spurious, although some of them may actually be correctly identified. However, nuclides like ^{123}I at CA002 and ^{137}Cs at FI001 were considered correct.

The reason for these assumptions was that peak explanation shares of each nuclide were not easily available in the PIDC database, which made the verification of nuclide associations difficult, especially in an automated investigation. Even with these simplifications, the comparison results should be reliable, *because the same criteria were used consistently for both software packages.*

The condensed results of the comparison are shown in Table 7.1. These direct comparison results gave a clear indication of SHAMAN's better peak identification percentage both in short-cycle and in long-cycle spectra. This advantage was compensated by a disadvantage, a larger number of spurious nuclides, but the difference in this respect was very modest, especially when the size difference of the respective gamma libraries is taken into account.

The average net difference in peak identification was slightly over 2 peaks per spectrum, one of which was significant and the other insignificant. This result coincided well with the results obtained by an experienced PIDC analyst, who manually compared the results of older versions of SHAMAN and Genie-NID in August–September 1997 [154]. The observed difference was approximately equal for both short-cycle and long-cycle spectra, but a strong variation from one station to another could be seen. This was probably due to the variation in operating procedures and equipment at different stations and the differences in natural radiation levels around the world, leading to varying identification challenges.

Station	Number of spectra	Shaman		Genie-NID	
		Additional peaks per spectrum not explained by Genie-NID	Number of spurious nuclides per spectrum	Additional peaks per spectrum not explained by Shaman	Number of spurious nuclides per spectrum
AR001 (short-c)	25	1.12 + 1.60	0.04	0.04	0.00
AU001 (short-c)	50	0.82 + 1.16	0.34	0.00	0.06
CA002 (short-c)	50	0.68 + 0.90	0.24	0.08	0.46
KW001 (short-c)	25	1.92 + 0.88	1.04	0.08	0.16
SE001 (short-c)	50	0.84 + 0.84	0.44	0.10	0.06
US001 (short-c)	50	1.12 + 0.88	0.12	0.06	0.02
total (short-c)	250	1.08 + 1.04	0.37	0.06	0.13
CA001 (long-c)	25	1.00 + 0.84	0.16	0.04	0.08
FI001 (long-c)	25	1.20 + 1.84	0.64	0.32	0.36
total (long-c)	50	1.10 + 1.34	0.40	0.18	0.22
total (all)	300	1.09 + 1.12	0.38	0.09	0.15

Table 7.1: Comparison between the identification results of Shaman and Genie-NID. Numbers of peaks and nuclides are averages in the corresponding spectrum set. Total averages were calculated with an equal weight for each station. The number of additional explained peaks by Shaman in the third column has been divided into significant (on the left) and insignificant (on the right) peaks.

The average net difference in the number of spurious nuclides was 0.23 nuclides per spectrum for this mixture of short-cycle and long-cycle spectra. There was an even stronger station-dependent variation in this number, but this was natural with so small a number of spurious nuclides. This difference was larger than that presented in Sec. 6.1, mainly due to a reduction in the number of spurious nuclides identified by Genie-NID. According to a PIDC radionuclide staff member, the average number of spurious nuclides in short-cycle spectra had been 0.29 during the first eight months of 1997 [140], but the version of November 1997 seemed to reach a substantially lower level, 0.13 in this test set as seen in Table 7.1. Changing a detector-specific library lookup tolerance to a spectrum-specific tolerance in the PIDC-software was credited with this improvement.

The statistical significance of the observed differences was proven by handling the randomly chosen 300 spectra as paired observations for the number of explained peaks, denoting the difference as ΔN_p, and the number of spurious nuclides, denoting the difference as ΔN_n. The observed distributions of these two random variables are shown in Fig. 7.1. For both variables, the value of Genie-NID was subtracted from that of Shaman.

According to a Chi-square goodness of fit test, the distributions of ΔN_p and ΔN_n did not represent a normal distribution very well. Therefore, the standard t-test was not applicable. In any case, a nonparametric statistical test could be applied to both distributions. The Wilcoxon signed-rank test for paired observations, a standard nonparametric method presented in textbooks on statistics (e.g., Ref. [155, Sec. 8.7 & 10.6]) was well suited for this situation. This method is used to test the hypothesis that the medians of two distributions,

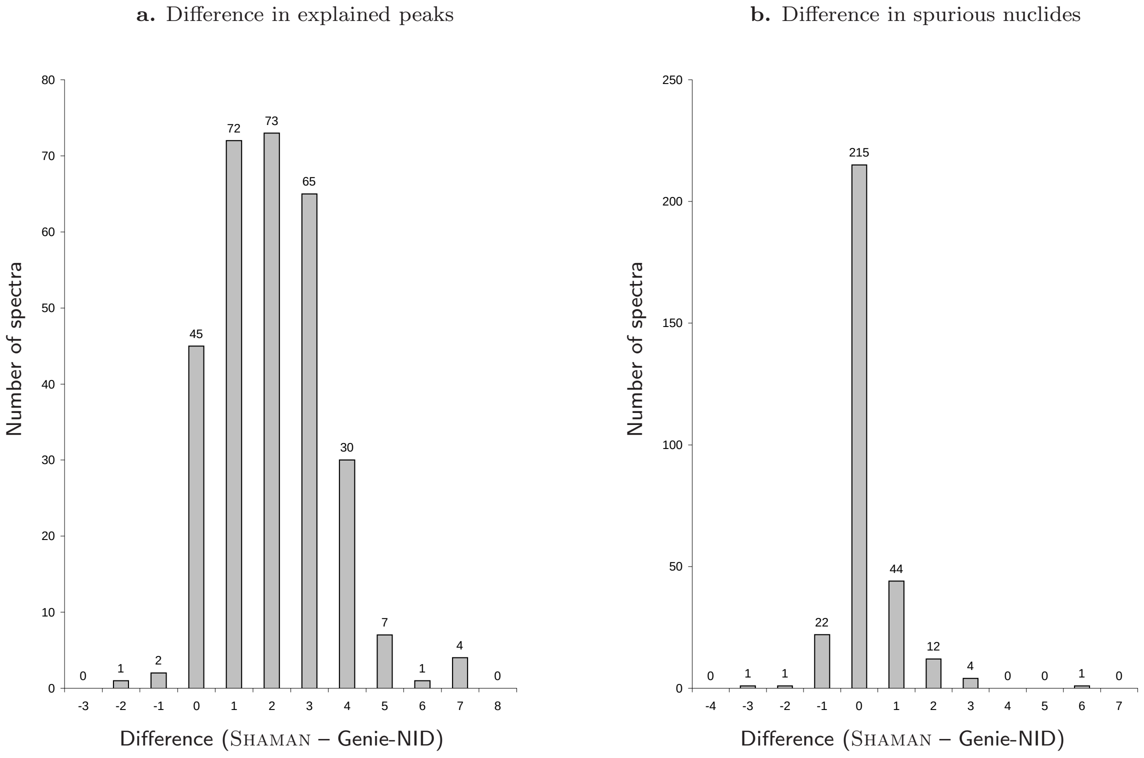

Figure 7.1: Difference in performance of Shaman and Genie-NID, shown by paired comparisons of 300 randomly selected PIDC spectra. **a.** The difference in the number of explained peaks per spectrum ΔN_p. The medians of the number of peaks explained by these two software differ at a >99.99 % confidence level. **b.** The difference in the number of spurious nuclides per spectrum ΔN_n. The medians of the number of spurious nuclides identified by these two software differ at a 99.95 % confidence level. (See discussion.)

observed as paired data, are equal and it does not require normally distributed variables.

In the Wilcoxon signed-rank test, the absolute values of n observations are ordered from smallest to largest and ranked from 1 to n. If ties occur, each member in a tied group is given the midrank of the group. The test statistic W is the smaller of the sum of ranks of negative observations and the sum of ranks of positive observations. One half of the ranks of zero-differences are added to the negative sum and one half to the positive sum. With $n > 50$, the statistic $(W - E[W])/\sqrt{\mathrm{Var}[W]}$ is approximately distributed as a standard normal variable with $E[W] = n(n+1)/4$ and $\mathrm{Var}[W] = n(n+1)(2n+1)/24$ [155]. Using the statistical test on the data presented in Fig. 7.1, the observed difference ΔN_p was proven to be above 0 at a >99.99 % confidence level, and the observed difference ΔN_n was proven to be above 0 at a 99.95 % confidence level.

The same statistical test was applied for the difference in the number of *significant* explained peaks. As shown in Table 7.1, the average difference was 1 significant peak per spectrum, for both short-cycle and long-cycle spectra, and using the Wilcoxon signed-rank test, this difference was proven to be above 0 at a >99.99 % confidence level. This was an indirect proof for the difference in peak identification percentages presented in Sec. 6.1 without a statistical proof. The difference in percentages in Sec. 6.1, $99.3\,\%-97.2\,\% = 2.1\,\%$, translated to 0.9 peaks in an average short-cycle spectrum with 41 significant peaks. This value is sufficiently close to the observed value 1.08 significant peaks in the direct comparison of Table 7.1.

The adopted concept of radionuclide operations at the CTBTO is that human analysts review the identification results given by the automatic analysis pipeline. They attempt to explain all unexplained spectrum peaks and to discard all spuriously identified nuclides. In this operational mode, a crude estimate for the effect of performance differences on analyst review time t_r can be calculated on the basis of the observed differences in peak identification and spurious nuclide identification. As a first approximation, we can assume a linear dependence of t_r on the the number of spurious nuclides N_n and the number of unexplained peaks $N_u = N - N_p$ where N is the total number of peaks found. The difference between single spectrum review times Δt_r for a system utilizing SHAMAN as its nuclide identifier and another one utilizing Genie-NID is in this approximation:

$$\Delta t_r = A \times \Delta N_n + B \times \Delta N_u = A \times \Delta N_n - B \times \Delta N_p \ ,$$

where the non-negative factors A and B are the analyst processing time per a spurious nuclide and an unexplained peak, respectively. If we use $A = 2\,\mathrm{min/nuclide}$ and $B = 2\,\mathrm{min/peak}$[a] and the observed performance differences $\Delta N_n = +0.24$ nuclides per spectrum and $\Delta N_p = +2.12$ peaks per spectrum, the difference would be $\Delta t_r \approx -4$ minutes per spectrum. When the number of stations in the complete IMS network is 80 and each of them sends one spectrum to be reviewed per day, the difference in the single spectrum review time would accumulate total savings of 5 efficient working hours per day when utilizing SHAMAN as the nuclide identifier instead of Genie-NID. With other values for A and B or with different concepts of operations, the daily savings in review time vary, but this calculation shows that they could be macroscopic.

7.1.2 Quantitative Differences in Identification Results

There are differences in quantitative identification results, i.e., nuclide activities, concentrations, and MDCs, between SHAMAN and Genie-NID. Generally, since activities can be interpreted as peak areas multiplied with a few constants, it would make more sense to compare peak areas instead of activities. In this case, however, SHAMAN used exactly the same peak data as Genie-NID, so the comparison of activities is motivated.

Due to time limitations, quantitative differences were not quantified in 1997, but the sources for differences were listed as follows:

1. **Differences in library values of emission probabilities.**

   The reference libraries of SHAMAN and Genie-NID had been compiled from different sources, which led to several differences discussed in Ref. [VI]. Any differences in emission probabilities always lead to differences in calculated activities.

---

[a] These are realistic estimates based on the average single spectrum review time t_r = 10–15 minutes observed at the PIDC.

Nuclide	Gamma energy (keV)	Emission probability (far geom.)	Emission probability (close geom.)	Correction coefficient
^{134}Cs	563.2	8.4	6.05	1.389
	569.3	15.4	11.12	1.385
	604.7	97.8	80.38	1.217
	795.8	85.4	70.16	1.217
	801.8	8.7	6.44	1.352
	1365.1	3.0	3.96	0.757
	1400.5	0.0	6.03	0.000
^{137}Cs	661.6	85.2	85.2	1.000
^{140}Ba	162.6	5.6	5.07	1.104
	304.8	3.4	2.79	1.217
	423.8	2.5	2.21	1.129
	537.3	20.0	17.71	1.129
^{140}La	328.8	20.0	14.66	1.364
	487.0	46.7	37.10	1.259
	815.7	22.8	21.88	1.042
	1596.4	96.0	79.56	1.207
^{212}Pb	115.2	0.592	0.487	1.217
	176.7	0.052	0.047	1.222
	238.6	43.3	43.29	1.000
	300.1	3.28	2.70	1.214
	415.2	0.143	0.222	0.643

Table 7.2: An illustration of the effect of coincidence summing on effective emission probabilities of some radionuclides. These calculations apply to a 40 % HPGe-detector in a close measuring geometry for a compressed air filter.

2. **Activity determination method: key-line activity versus least-squares activity.**

   Shaman calculates all nuclide activities at the same time using the SVD-algorithm [5, p. 58–63]. This method yields the best activity estimates in a least-squares sense, dividing areas of the spectrum peaks to each associated nuclide (so-called "interference correction"). Genie-NID, on the other hand, calculates activities based on a pre-selected key line. This gives the same activity as Shaman's method for single-line nuclides which are associated to their peaks alone, but the values for all other nuclides may differ substantially. There are options for different activity calculation methods in Genie-NID and Shaman can report a key-line activity, but in 1997 the selected methods differed as presented.

3. **Coincidence summing correction.**

   Shaman corrects activities for coincidence summing effects, which are strong in spectra measured in a close geometry. Coincidence summing modifies peak area ratios from the ones in a nuclide library by reducing the sizes of cascade peaks and increasing the sizes of crossover peaks. It also produces coincidence sum peaks (see Sec. 4.7) that would not be detected in a non-summing measurement setup (far geometry).The significance of this effect is illustrated in Table 7.2, where coincidence summing corrections calculated with Shaman's method are shown for a typical 40 % HPGe-detector in an air filter measuring geometry. Large correction coefficients differences can be seen between different nuclides and between gamma lines of a single nuclide.

The effects of coincidence summing vary from one nuclide to another, but in a typical case the calculated nuclide activities will be too low if the summing effect is not compensated for. Hence, calculating coincidence corrections, implemented in Shaman but not in Genie-NID, yields more accurate activity estimates and quantitative explanations for coincidence sum peaks.

4. **Self-absorption correction.**

   A correction for self-absorption is important if the size or composition of the efficiency calibration sources differ significantly from the actual samples. In a typical IMS station, self-absorption in the actual samples is similar to that in the calibration samples and no correction is needed. However, if any additional absorption is not compensated for, the calculated nuclide activities will be too low. In Shaman and the PIDC-software, methods to compensate for self-absorption effects have been implemented. If these corrections are used, they will introduce additional differences between activities calculated by Shaman and the PIDC-software, because the correction methods are different.

5. **Baseline integration interval in MDC calculation.**

   All the phenomena above have an effect on Minimum Detectable Activity (MDA) or Concentration (MDC) calculations. However, there is another reason for the differences between the MDC-values that has major importance: the difference in the baseline integration interval. Both Shaman and Genie-NID apply the basic method of Currie [94] at the 95 % confidence level, but the choice of the baseline integration interval is very different. Shaman uses an integration interval of $[-1.4\sigma, +1.4\sigma]$, where σ is the width parameter of a Gaussian peak given by the peak shape calibration. This interval has been derived to be the optimum in the signal-to-noise sense [5, p. 41–43]. Genie-NID, on the other hand, uses an approximately 4 times larger integration interval. Since the calculated MDC is roughly proportional to the square root of the integrated baseline, the MDCs of Genie-NID are typically a factor of 2 larger than those calculated by Shaman. Therefore, the MDCs calculated by these two software are not comparable.

The first four phenomena lead to differences in activities and concentrations calculated by Shaman and Genie-NID. Since these differences are systematic, they do not affect monitoring of concentrations of single nuclides, which is done for sample categorization purposes, provided that the same software is constantly used. However, no concentration or MDC values should be compared across software packages without great caution.

However, the most important effect of quantitative identification differences is that radionuclide activity ratios, which can be utilized in backtracking calculations in a possible release situation, may become distorted, leading to a large error in release timing. The importance of coincidence and self-absorption corrections is especially emphasized in activity ratio studies. Coincidence correction is important if one of the nuclides experiences stronger summing effects than the other. This is the case with the following, potentially interesting nuclide pairs: ^{99}Mo/^{99m}Tc, ^{134}Cs/^{137}Cs, and ^{140}Ba/^{140}La (see Table 7.2). Similarly, an accurate self-absorption correction is necessary for nuclide pairs, where the primary gamma energies are far apart.

7.2 Comparison between SHAMAN and the IDC Nuclide Identification

7.2.1 Evaluation Methods and Material

As mentioned in Sec. 2.2.4, the Finnish NDC (FiNDC) has been running a UNISAMPO–SHAMAN analysis pipeline since July 1999. In order to quantify the FiNDC performance, a comprehensive study was carried out in 2001 where the FiNDC pipeline results for IMS spectra were compared to those obtained by the International Data Center (IDC). A complete presentation of the study is available in Ref. [XII], but the most essential findings are presented here in a manner similar to Ref. [XIV].

The subject of the evaluation was the pipeline analysis, i.e., results obtainable without user intervention. Therefore, the NDC pipeline results were compared to those obtained by the IDC analysis pipeline. The automated results can naturally be improved in an interactive session by a human analyst, as is done constantly at the IDC and selectively at the FiNDC. The quality of the reviewed final result is very much dependent on the analyst's experience and the time and software resources available. Given the same review resources, however, a higher final quality can be reached when the automated pipeline results are of higher quality.

Since the IDC analysis pipeline produces a report called Automated Analysis Report (ARR) in a fixed format, we developed a similar reporting format for SHAMAN. The information in the ARR is quite limited, which reduced the possibilities for comparisons. For example, the ARR format does not contain peak explanation shares, i.e., which part of the peak is explained by the nuclides associated with it. Thus, each peak that had any degree of explanation by an associated nuclide had to be counted as completely explained. We made additions such as the peak share and peak significance to our ARR report format and recommended the same upgrade to the CTBTO because we have found these parameters useful in interpreting spectrum analysis results.

The IMS air filter sample spectra and the corresponding ARR reports are received by the FiNDC through the IDC's data subscription mechanism. After processing the spectra in the pipeline, the NDC directory tree contains ARR's of both the IDC and the NDC. Scripts and programs were implemented to match the ARR-pairs and to make different comparisons. The following parameters were compared:

– peak areas,
– peak area uncertainties,
– numbers of found peaks,
– numbers of explained peaks,
– numbers of unexplained peaks,

– peak identification percentages (how many peaks were explained),
– differences in nuclide identifications, and
– numbers of spurious (incorrect) identifications.

At the time of the study, there were eight radionuclide stations sending data to the IDC on a daily basis. Five of these stations were certified IMS stations, three had a test station status. All air filter spectra from these stations with an 24/24/24-hour operation cycle and with acquisition start between August 2000 and March 2001 were used for the investigations. This led to a test set of 1 428–1 518 spectra, depending on the exact date of each investigation. Only the spectra which had severe quality problems, e.g., a large temporary gain shift, or whose IDC analysis report was not available at the FiNDC were disqualified from the comparisons.

7.2.2 Evaluation Results

The peak areas determined by the FiNDC and IDC analysis pipelines showed some discrepancies especially for small peaks. This is to be expected, since UniSampo uses a fixed-width area determination, whereas the analysis pipeline of the IDC is mainly based on Canberra's Genie-2000 Spectroscopy Tools software package and uses its free-width peak fitting option [156]. Also the number of channels used for baseline determination is different in the FiNDC and IDC pipelines. Even the areas of large peaks showed a difference of the order of 1 %. It was not possible for us to examine the reason for this, but the difference between the peak functions (lower and upper tail components used at the FiNDC, not at the IDC) is a probable explanation.

The peak area uncertainties showed a large discrepancy. The discrepancy became apparent when observing the largest peaks found in the spectra for which the peak area uncertainty should tend toward the Poissonian uncertainty or $\Delta A = \sqrt{A}$. The FiNDC peak data clearly obeyed this relation as seen in Fig. 7.2, but in the IDC peak results, the peak area uncertainties seemed to have been multiplied with a factor of 1.6. This could be due to a confidence level different from the 1σ level used by the FiNDC, but no confirmation for this could be found in the documentation available to us.[b]

The peak search algorithms of UniSampo and the IDC software are closely related. However, the FiNDC applies a lower peak search threshold than the IDC pipeline, complemented with a stricter peak significance test, which together lead to a larger number of peaks found and simultaneously a smaller number of spurious peaks. These facts are clearly visible in Fig. 7.3. In the test set of 1 518 spectra, the FiNDC pipeline found 4.2 peaks more per spectrum on the average (9 % increase). Simultaneously, the FiNDC pipeline explained 5.6 more peaks per spectrum than the IDC pipeline (14 % increase). In other words, the number of unexplained peaks was 1.4 peaks or 40 % less in the FiNDC results for an average spectrum.

The latter result is a combination of the differences in peak search and the sophistication of nuclide identification methods. The IDC pipeline utilizes the identification method of Genie-2000 with a tailored library of only about 120 weapons-related and natural radionuclides,

---

[b] The constant difference factor of 1.6 disappeared in the first IDC software update after the publication of Ref. [XII], hinting to a confidence level difference.

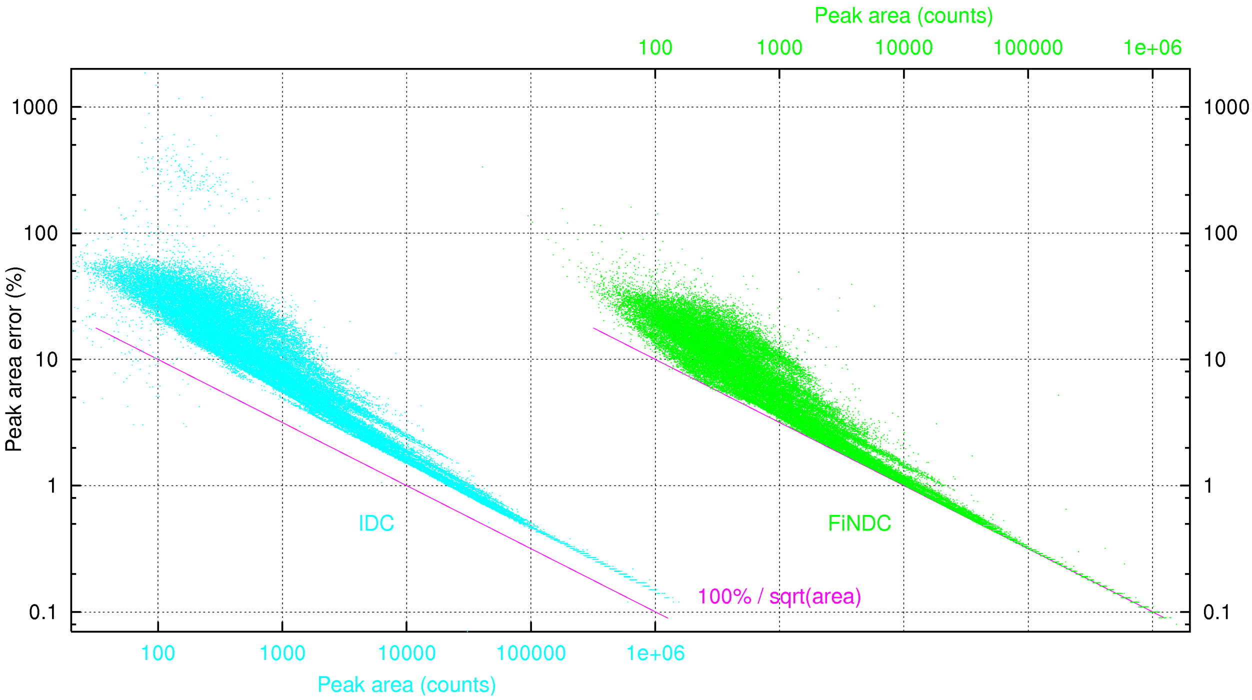

Figure 7.2: Peak area uncertainties as a function of peak area for a set of 52 744 peaks found by both IDC (cyan points) and FiNDC (green points) in 1 518 IMS spectra. The red lines show the Poissonian uncertainty. The data points have been shifted horizontally for clarity.

whereas at the FiNDC nuclides are identified with SHAMAN and its comprehensive library. In addition, SHAMAN features explanation capabilities for escape peaks, X-ray escape peaks, and sum peaks (see Ch. 4), all of which are seen in IMS spectra. Thus, the performance difference is understandable.

Each peak that receives any degree of explanation by the identified nuclides is classified as explained in the ARR reports of the IDC. The ratio of the explained peaks to all found peaks is referred to as the “peak identification percentage” P'. These percentages are rounded to the closest integer value. The distributions of peak identification percentages of the FiNDC and IDC pipeline are plotted in Fig. 7.4. The distributions are clearly different. The median values of the FiNDC and the IDC were 96.7 % and 92.3 %, respectively.

Since there were virtually no anthropogenic nuclides to be identified in the test spectra, the peaks additionally found and explained by the FiNDC belonged to natural nuclides. However, there is no reason to assume that the performance difference would be inverted if peaks of activation and fission products were present in spectra. The FiNDC pipeline finds more peaks than the IDC pipeline with the settings in use and the performance of its identification part has been shown to remain on a high level even for air filter spectra with anthropogenic contents (see Sec. 6.2 and 6.3).

In addition to the figures above, estimates for the numbers of spuriously identified nuclides and peaks are evidently required. Strictly speaking, this would require a careful manual analysis or a set of synthesized spectra with exactly known nuclide and peak contents. However, we developed a simplistic strategy for the assessment of spurious nuclide identifications: since

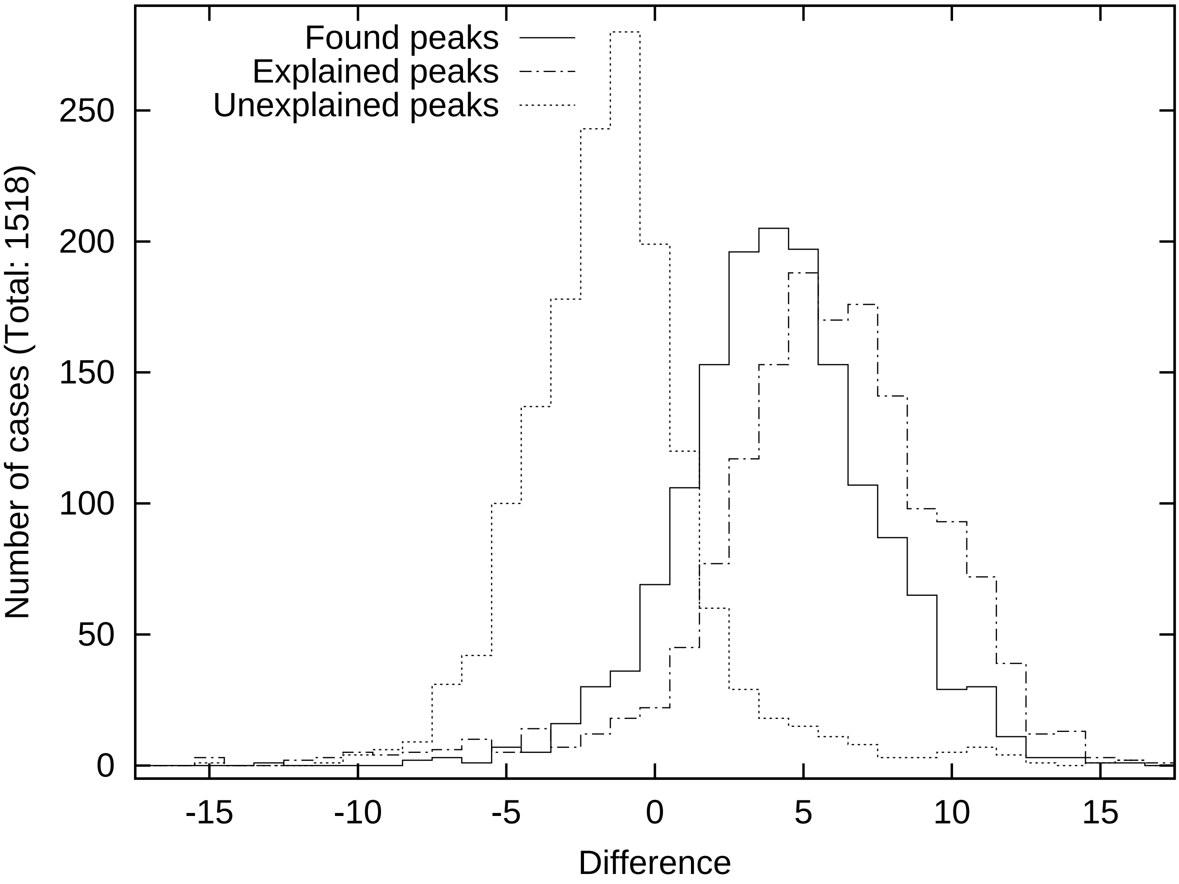

Figure 7.3: The distributions of the differences (FiNDC − IDC) of the number of found, explained and unexplained peaks in a test set of 1 518 spectra.

all spectra in the test set were air filter spectra without anomalous contents, all anthropogenic nuclides in the pipeline reports were declared spurious. Only the following exceptions that were present in some samples were accepted as correctly identified: ^{60}Co, ^{99m}Tc, ^{131}I, and ^{137}Cs. This method does not give the exact truth, but the spurious identification estimate is realistic and it treats both pipelines equally.

The distribution of the difference between the number of spurious nuclide identifications as defined above is shown in Fig. 7.5. It can be seen that the FiNDC results typically contained more spurious nuclides than the IDC results. On the average, there were 0.7 more spurious nuclides in the FiNDC results, but only a fraction of them are considered relevant by the CTBTO [146]. The difference in CTBTO relevant spurious nuclides was only 0.05 per spectrum on the average. A difference in this direction is to be expected when the sizes of the reference libraries and the peak search thresholds are considered. The difference is, however, quite acceptable since it is much easier for a human analyst to discard spurious identifications than to find explanations for unexplained peaks. Moreover, a comprehensive library is fundamental to nuclide identification in an eventual fresh release situation.

The findings presented in this section are not directly comparable to those presented in Sec. 7.1, because here two completely separate analysis pipelines are compared while only

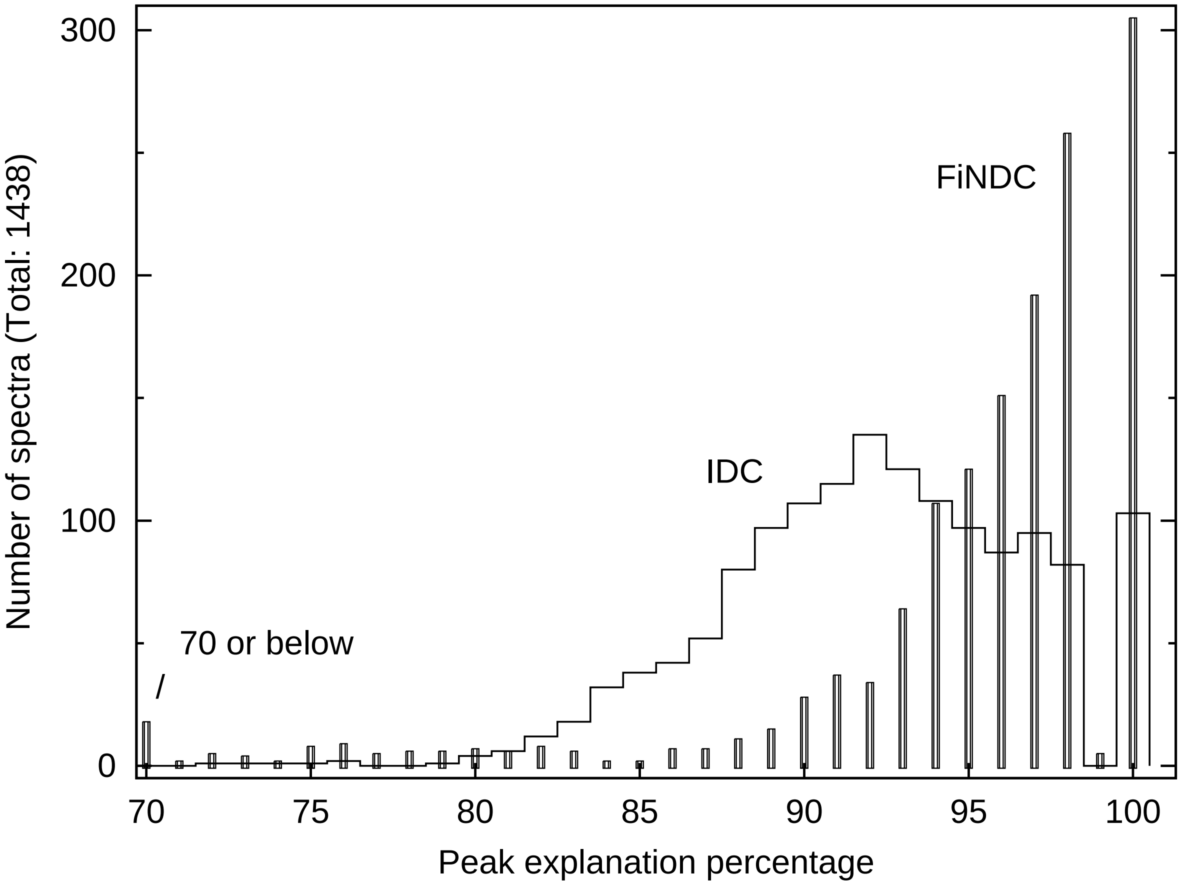

Figure 7.4: The distributions of the peak identification percentages in a test set of 1 438 spectra. The reported percentages are rounded integers, so the value 99 % is possible only for spectra with at least 67 peaks.

the nuclide identification component differed in the previous comparison. Nevertheless, both comparisons show that SHAMAN's identification results are significantly better than those obtained with the SAMPO-method implemented in Genie. A comparison between identification results obtained with SAMPO or UNISAMPO and SHAMAN has not been accomplished, but according to some test analyses the results obtained in these SHAMAN–Genie-comparisons can most probably be generalized.

All comparisons concentrated on results obtained without user intervention. The pipeline results can naturally be improved with interactive analysis, but in SHAMAN's case, the need for improvements during the review is clearly reduced. However, SHAMAN's graphical user interface is well applicable for closer investigations: usually the analysis results for ordinary air filter spectra are fully acceptable after one or two interactive iterations.

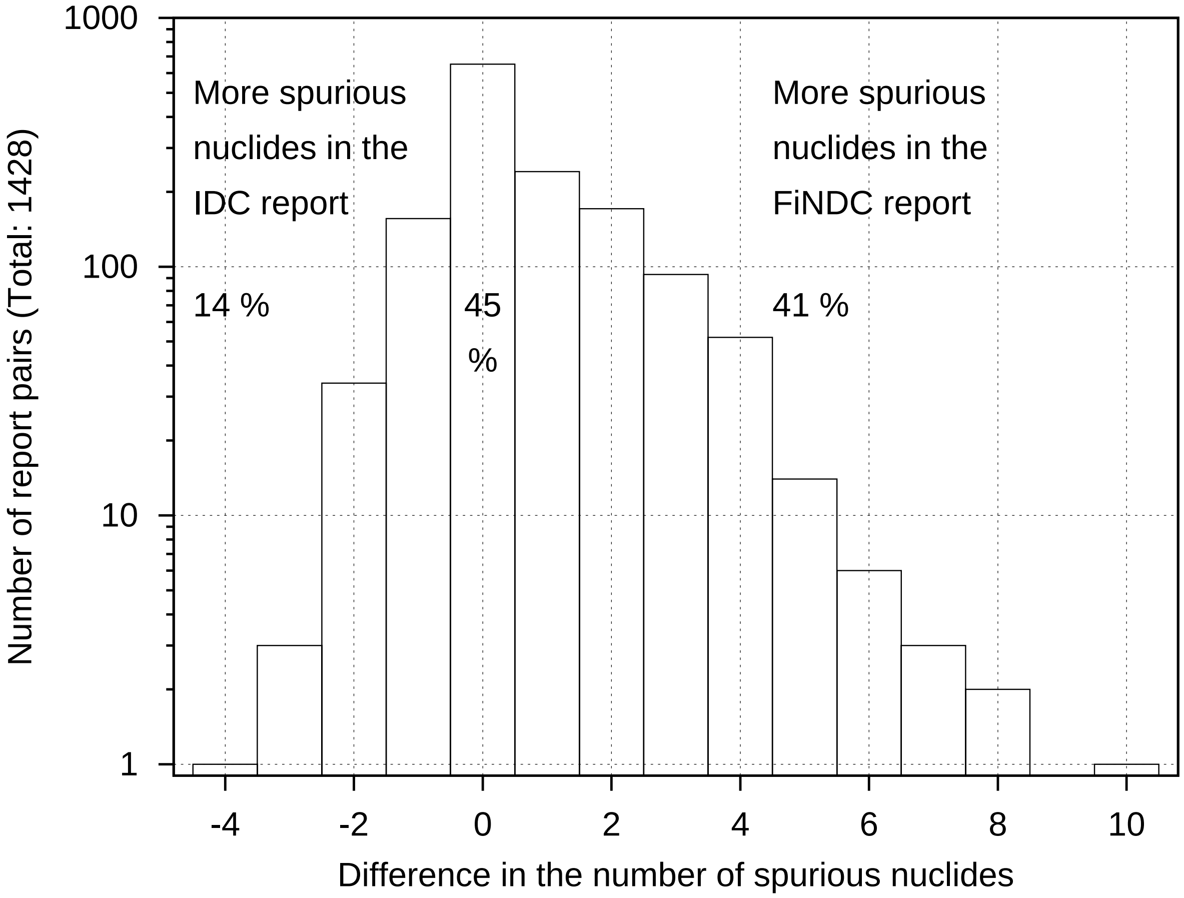

Figure 7.5: The distribution of the difference between the number of spurious nuclide identifications in a test set of 1 428 spectra.

Chapter 8

Summary and Conclusions

High-resolution gamma-ray spectrometry is an analysis method well suitable for monitoring airborne radioactivity. Many of the natural radionuclides as well as a majority of anthropogenic nuclides are prominent gamma-ray emitters. There are a few exceptions, most notably ^{3}H, ^{14}C, ^{90}Sr, and ^{239,240,241}Pu that are detectable through their α or β-radiation only. Therefore, gamma-ray spectrometry must be complemented with other methods for a full monitoring coverage, but it is generally the principal analysis method. With gamma-ray spectrometry different radionuclides are readily observed at minute concentrations that are far from health hazards. Thus, measurements during normal conditions and minor anthropogenic releases can be utilized to increase sensitivity of analysis methods and to learn to distinguish between source terms from different release types.

The gamma-ray spectrometric analyses applied in air monitoring programmes can be divided into particulate measurements and gas measurements, of which the former category is technologically more mature. In this work, methods applicable for particulate sample analysis have been presented, implemented in analysis software, and evaluated with a wide variety of cases. Our goal has been to develop a collection of tools that enables a complete quantitative explanation of all components of a measured gamma-ray spectrum with a minimum of user intervention.

In this work, a comprehensive assessment of analysis methods for airborne radioactivity was performed and substantial improvents in an analysis software package, expert system Shaman, were developed and demonstrated. The latter included:

- tools for the correct identification and estimation for small interfering peaks, e.g., escape peaks and sum peaks,
- adapting, correcting, extending and integrating a comprehensive nuclear data library derived from the ENSDF data compilation,
- extensive benchmarking and testing of Shaman, comparing it to earlier implementations and other analysis systems.

The most essential findings are summarized in the following.

The importance of accurate and precise calibrations cannot be overestimated. Energy and peak efficiency calibrations are fundamental for spectrum analysis, since they are needed to convert the peak positions and areas to energies and emission rates and further to nuclide identifications and quantifications. Use of a shape calibration can also be recommended for its stabilizing effect on peak search and area determination. If measurements are made in a close geometry, a total efficiency calibration is also needed in order to enable corrections for true coincidence summing effects.

Sufficiently accurate energy and shape calibrations can be derived internally, i.e., from the spectrum being analyzed. The internal calibrations are based on the nominal ones whose accuracy gradually decreases. If the drift is too severe, the internal calibration methods may fail, but a failure in calibration update, utilizing the simple methods presented in Sec. 4.3, usually indicates some kind of hardware problems.

Peak and total efficiency calibrations cannot be made internally, since they are based on known activities of calibration sources. Typically, the accuracy of efficiency calibration dictates the best accuracy that can be obtained for nuclide activities. In the spectrum analysis methods for escape peaks and true coincidence correction, efficiencies are needed for a wide energy range, typically 15–4000 keV that is wider than the energy range covered by the spectrum, so the extrapolation properties of the calibration curves need to be examined. Especially the low-energy end is a challenge, since low-energy gamma-ray sources with a sufficiently accurate emission probability and reasonable half-life are scarce.

The accuracy requirements for total efficiency are not as high as for peak efficiency, but a calibration is in any case needed if using a close geometry. An empirical total efficiency is naturally preferred, but there are three alternatives for it: utilizing peak efficiency and a parameterized total-to-peak ratio, or calculating total efficiency from a phenomenological approximation, or modeling the source-detector geometry and utilizing Monte Carlo simulation. The obtainable accuracy increases in the given order, but so does the effort. The parameterized total-to-peak ratio has been utilized in our analyses and it seems to be applicable in routine monitoring analyses.

Since this work concentrates on spectrum interpretation, not peak analysis, the peak finding and quantification methods have not been investigated in detail. However, the critical limit for peak acceptance is important for identification purposes, so different methods for decision limit calculation were collected from literature and compared to the method implemented in our software Shaman. It was found that there is a method defined in a DIN standard that could replace the current method. The fundamental difference between these two methods is the choice of baseline summing interval: whether it is under an assumed peak or on both sides of it. A constant-factor type difference between these two methods was also observed. It is more or less a scaling issue that can be settled with a set of synthesized spectra.

In a high-resolution spectrum, all essential information is contained in the peaks. Most of them are full-energy peaks that can be explained with gamma-ray lines in a reference library. Among the full-energy peaks, however, there are annihilation escape peaks, X-ray escape peaks, coincidence sum peaks, and random sum peaks that are not tabulated. They need to be explained quantitatively if a complete interpretation of a measured spectrum is aimed at. Calculation methods for these special peaks are presented, implemented, and evaluated in Secs. 4.5–4.8 of this work.

The method of Heydorn and Rhee for estimating annihilation escape peak areas was found to result in peak explanation levels in the range 70 %–130 %. This range is acceptable for nuclide identification purposes, especially taking into account that the escape peaks are typically small. Thus, the method is applied in SHAMAN by default.

Based on an analogy with coincidence measurements, a simple method for random sum peak area prediction was developed in this work. We found that when a spectrum-specific constant was used as the essential parameter of the model, the method did not predict random sum peak areas very well. A better agreement was achieved when the energy difference of the summing gamma-rays was factored into the parameter. With the improved model, the estimated peak areas were already within a factor of 2 of the observed ones. This accuracy may be sufficient in some applications, but the model needs to be refined before it can be used routinely. Additionally, its applicability with different MCA hardware needs to be investigated.

We have extended the traditional coincidence correction method by Andreev et al. to calculate the areas of true coincidence sum peaks. The sum lines are added to the decay schemes of interesting radionuclides with a small initial emission probability that is subtracted from the coincidence corrected emission probability afterwards. The evidence presented in this work shows that the method works accurately and thus, it is applied in SHAMAN by default. There are two omissions in the calculation model of SHAMAN: coincidence summing of gamma-rays from non-consecutive states and sum peaks of X-rays are not accounted for. The calculation method for the former is presented in this work and for the latter in literature.

We also extended the calculation method for X-ray escape peaks above the traditionally applied 100 keV limit. With current Ge detectors, X-ray escape peaks may be observed at all energies. Above 100 keV, the escape-peak-to-photopeak ratio is of the order of 10^{-4}–10^{-3}, so the escape peaks are seen only beside the largest photopeaks. An important case in air filter spectra is the 228 keV escape peak below the prominent 238 keV peak of ^{212}Pb, because its alternative explanation would be ^{140}Ba, a short-lived fission product. It was shown that the accuracy of our method is acceptable for identification purposes. The method seems to fail if there are charge-collection problems in the Ge detector, but assuming a healthy detector, the calculation is used in SHAMAN by default.

Expert system SHAMAN combines a comprehensive ENSDF-based reference library with an inference engine that applies pruning rules to select the acceptable candidate nuclides, and with a collection of calculational methods including the above-mentioned ones. Its performance with particulate air filter spectra has been evaluated in this work with four different spectrum sets. Two of the sets have been produced at different times by a global monitoring network that has been set up for the Comprehensive Nuclear-Test-Ban Treaty, and the two other sets, measured after two minor anthropogenic releases, by the Finnish monitoring network run by STUK.

In the first CTBT-related set of 250 spectra the peak analysis was made with tailor-made analysis software based on Genie. Its output was analyzed with two SHAMAN versions and they both obtained a nuclide identification percentage and a peak identification percentage of approximately 99 % each. The spurious nuclide percentage was 4–7 %, translating to 0.4–0.7 spurious identifications per spectrum. These results were obtained in a completely automated analysis mode, so they can be considered excellent.

The second CTBT-related set included 6 161 spectra that were not analyzed manually unlike the first spectrum set, only with an automated UniSampo–Shaman pipeline system. Without a manual analysis we presented uncorrected peak identification percentages and uncorrected average numbers of spurious nuclides of 96.4 % and 0.96, respectively. Both figures are less ideal than for the previous spectrum set, but the explanation is that a smaller peak search threshold has been utilized in the peak analysis phase. Therefore, the average number of peaks is much larger than that in the first set and especially the share of small peaks has increased. The performance level can still be considered very good, even if there is some room for improvement.

The spectrum sets of the national network had been measured after minor releases from Novaya Zemlya in August 1987 and Sosnovyy Bor in March 1992. Both spectrum sets were analyzed with Sampo in interactive mode and then identified with Shaman, so the results are not directly comparable to the automated analysis results referred to above. However, these two spectrum sets showed that the identification performance of Shaman is not degraded in cases where air filter spectra contain fission and activation products in addition to natural radioactivity.

In the Novaya Zemlya spectrum set the number of anthropogenic nuclides was only 3.6 per spectrum and the observed peak and nuclide identification percentages practically the same as in routine air filter spectra. Also the number of spurious nuclides remained at 1.0 per spectrum. In the Sosnovyy Bor spectrum set there were 6.0 anthropogenic nuclides per spectrum on the average and the set included the most complicated spectra in these spectrum sets, with over 30 present nuclides and over 80 peaks in 7 spectra. This is why Shaman's performance was slightly degraded: nuclide and peak identification percentages were 96.5 % and 96.6 %, respectively, and the number of spurious identifications 1.9 per spectrum on the average. However, the degradation cannot be considered severe and with some assistance from the human analyst, Shaman can reach a high level.

The identified nuclide contents of the Novaya Zemlya and Sosnovyy Bor spectra did not contain any big surprises, but some interesting observations were made. In addition to the expected fission products, the activation products ^{122}Sb and ^{134}Cs were identified in the Novaya Zemlya spectra. Antimony is known to be used in nuclear devices, so the ^{122}Sb observation cannot be ruled out, although it was made in one spectrum only. On the other hand, the activity of ^{134}Cs relative to ^{137}Cs matched with Chernobyl resuspension in all spectra where it was identified, so it was concluded that the source of ^{134}Cs was probably not the Novaya Zemlya event.

The Sosnovyy Bor spectra were likely to contain more activation products among fission products, since the source term came from a nuclear reactor. This was also the case: ^{122}Sb, ^{134}Cs, ^{136}Cs, ^{237}U, ^{238}Np, and ^{239}Np were identified in these spectra. In some cases, the ^{134}Cs activity relative to ^{137}Cs matched again with Chernobyl resuspension, but in 13 of 24 observations the ratio was clearly different. The heavy activation products ^{237}U and ^{239}Np are also produced in nuclear devices that contain ^{238}U. In contrast, the production of ^{238}Np is less probable, since it is produced by neutron capture from ^{237}Np that is produced in the decay of ^{237}U. ^{237}Np is also produced by the α-decay of ^{241}Am but its half-life is much longer. Thus, we conclude that identification of ^{238}Np is more typical of a reactor release than of a bomb release.

The performance of SHAMAN was benchmarked against another identification software based on Genie in two intercomparisons. In both cases, the results from automated analysis were compared, thus neglecting the effect of a human analyst that can compensate for errors in automated results. The first intercomparison truly measured the difference in nuclide identification, since the same peak results were fed into both identification implementations. This direct comparison showed that SHAMAN provided an identification for two more peaks in an average spectrum, while it identified 0.23 spurious nuclides more than the competitor. We proved these differences to be statistically significant and showed that the difference could mean macroscopic savings in the review time used by a human analyst. In our opinion, it is always easier for the human analyst to discard spurious nuclides than find explanations for unidentified peaks, so the settings of SHAMAN's discarding rules have been made conservative on purpose.

In the second intercomparison two automated analysis pipelines were compared where also the peak analysis results are different due to different software. Utilizing sets of over 1 400 spectra, our UNISAMPO–SHAMAN pipeline system was shown to find 4.2 peaks more per spectrum than the competitor. The difference in explained peaks was even larger, 5.6 peaks per spectrum, leading to a number of unexplained peaks that was 40 % smaller than that of the other software. This difference was explained with the peak search method of UNISAMPO and the sophistication of nuclide identification in SHAMAN. The figure-of-merit utilized by the CTBTO, uncorrected peak identification percentage, reflects the difference: the median values were 96.7 % for UNISAMPO–SHAMAN and 92.3 % for the other software. The disadvantage of our system was again a larger number of spurious nuclides: the difference was 0.7 nuclides in an average spectrum. This price is reasonable for the ability to use a comprehensive reference library that is fundamental to nuclide identification in a release situation.

In summary, the UNISAMPO–SHAMAN pipeline system is well applicable at any organization where environmental radioactivity is monitored. The system can also be utilized in other applications, but the current rule base of expert system SHAMAN has been tailored most comprehensively for air filter spectra. The analysis capabilities of the system can be improved in other applications of gamma-ray spectrometry by tuning the analysis parameters of the system. This should not require but a modest effort of an end-user who has the application-specific expertise and has some scripting experience since a major part of SHAMAN's "expertise" or "intelligence" is in its operating scripts and macros. Naturally, there is still room for improvement in the methodology of SHAMAN, but with the currently implemented features it is already a reliable expert system for nuclide identification.

More challenges are to be expected if more advanced measurement setups are used than a single Ge-spectrometer. Additional detectors can be utilized as an active Compton-suppression shield, lowering the background, or in a gamma-gamma coincidence measurement setup that could help to resolve some single-line and multiline nuclides from each other. Some of the current rules of SHAMAN would become useless when analyzing spectra measured in such setups and some new ones would probably need to be implemented. It is even possible that SHAMAN's reference library would need to be complemented with decay scheme details in these analyses. Nonetheless, a survey of applications that could benefit from the comprehensive nuclide identification philosophy of SHAMAN could be worthwhile at this stage.

References

Author's Own Publications

[I] J.J. Ala-Heikkilä, *Asiantuntijajärjestelmä* SHAMAN*in nukliditietokannan päivitys.* Raportti TKK-F-C152, Otaniemi 1993 (in Finnish).
Title in English: *Updating the Nuclide Reference Library of Expert System* SHAMAN.

[II] J.J. Ala-Heikkilä, *Coincidence Correction in* SHAMAN *for Environmental Measurements.* Report TKK-F-C166, Otaniemi 1994.

[III] P.A. Aarnio, J.J. Ala-Heikkilä, T.T. Hakulinen, J.T. Routti, "Expert System for Nuclide Identification in Gamma Spectrum Analysis". Journal of Radioanalytical and Nuclear Chemistry **193** 2 (1995) 219–227.

[IV] J.J. Ala-Heikkilä, *Expert System for Nuclide Identification in Environmental Gamma Spectra.* Report TKK-F-B159, Otaniemi 1995.

[V] J.J. Ala-Heikkilä, *Advance Prediction of Escape and Sum Peak Areas in Gamma-Ray Spectra.* Report TKK-F-B162, Otaniemi 1996.

[VI] J.J. Ala-Heikkilä, T.T. Hakulinen, P.A. Aarnio, H.I.K Toivonen, L.R. Mason, W.C. Evans, J.B. Bohlin, *Evaluation of Expert System* SHAMAN *in Processing Gamma-Ray Spectra at the Comprehensive Test Ban Treaty Prototype International Data Center — Phase I.* CMR Technical Report CMR–97/05, Arlington 1997.

[VII] P.A. Aarnio, J.J. Ala-Heikkilä, T.T. Hakulinen, P.A. Aarnio, *Evaluation of Expert System* SHAMAN *in Processing Gamma-Ray Spectra at the Comprehensive Test Ban Treaty Prototype International Data Center — Phase II.* CMR Technical Report CMR–97/13, Arlington 1997.

[VIII] J.J. Ala-Heikkilä, T.T. Hakulinen, P.A. Aarnio, M.T. Nikkinen, H.I.K. Toivonen, *Evaluation of Expert System* SHAMAN *in Processing Gamma-Ray Spectra at the Comprehensive Test Ban Treaty Prototype International Data Center.* Report TKK-F-B171, Otaniemi 1997.

[IX] P.A. Aarnio, J.J. Ala-Heikkilä, T.T. Hakulinen, M.T. Nikkinen, "Application of the Nuclide Identification System SHAMAN in Monitoring the Comprehensive Test Ban Treaty". Journal of Radioanalytical and Nuclear Chemistry **235** 1–2 (1998) 95–103.

[X] P.A. Aarnio, J.J. Ala-Heikkilä, T.T. Hakulinen, M.T. Nikkinen, "Expert System Based Radionuclide Identification". Pp. 36–50 in *Software for nuclear spectrometry.* IAEA-TECDOC-1049, IAEA Vienna 1998.

[XI] P.A. Aarnio, J.J. Ala-Heikkilä, T.T. Hakulinen, M.T. Nikkinen, "The Nuclide Identification System SHAMAN in the Verification of the Comprehensive Nuclear-Test-Ban Treaty". Journal of Radioanalytical and Nuclear Chemistry **248** 3 (2001) 587–593.

[XII] T.K. Ansaranta, J.J. Ala-Heikkilä, P.A. Aarnio, *Comparison of Radionuclide Data Analysis Results of the CTBTO/IDC and the Finnish NDC.* STUK (Radiation and Nuclear Safety Authority), Report STUK-YTO-TR 180, Helsinki 2001.

[XIII] P.A. Aarnio, J.J. Ala-Heikkilä, T.T. Hakulinen, M.T. Nikkinen, "Expert System SHAMAN and Verification of the Comprehensive Nuclear Test Ban Treaty". Pp. 93–101 in *Specialized software utilities for gamma ray spectrometry.* IAEA-TECDOC-1275, IAEA Vienna 2002.

[XIV] P.A. Aarnio, J.J. Ala-Heikkilä, T.K. Ansaranta, T.T. Hakulinen, "Analysis Pipeline for Air Filter Gamma-Ray Spectra from the CTBT Verification Network". Journal of Radioanalytical and Nuclear Chemistry **263** 1 (2005) 251–257.

[XV] P.A. Aarnio, J.J. Ala-Heikkilä, T.T. Hakulinen, M.P. Huhtinen, "High-Energy Proton Irradiation and Induced Radioactivity Analysis for Some Construction Materials for the CERN LHC". Journal of Radioanalytical and Nuclear Chemistry **264** 1 (2005) 51–60.

[XVI] P.A. Aarnio, J.J. Ala-Heikkilä, T.T. Hakulinen, M.T. Nikkinen, J.T. Routti, "WWW-Based Remote Analysis Framework for UNISAMPO and SHAMAN Analysis Software". Journal of Radioanalytical and Nuclear Chemistry **264** 1 (2005) 255–258.

[XVII] P.A. Aarnio, J.J. Ala-Heikkilä, T.T. Hakulinen, "Performance of UNISAMPO–SHAMAN with Gamma-Ray Spectra Containing Known Traces of Fission Products". Journal of Radioanalytical and Nuclear Chemistry **276** 2 (2008) 455–460.

[XVIII] P.A. Aarnio, J.J. Ala-Heikkilä, A. Isolankila, A. Kuusi, M. Moring, M. Nikkinen, T. Siiskonen, H. Toivonen, K. Ungar, W. Zhang, "*Linssi* — Database for Gamma-Ray Spectrometry". Journal of Radioanalytical and Nuclear Chemistry **276** 3 (2008) 631–637.

[XIX] P.A. Aarnio, J.J. Ala-Heikkilä, T.T. Hakulinen, "A Multi-User Server-Based Framework for SHAMAN Gamma-Ray Spectrum Identification Software". Journal of Radioanalytical and Nuclear Chemistry **276** 3 (2008) 651–655.

[XX] SHAMAN — *Expert System for Radionuclide Identification, Version 1.16.* User's Guide Version 1.9. Baryon Oy, Ltd. Espoo 2007.

Other SHAMAN Publications

[1] P.A. Aarnio, T.T. Hakulinen, *Asiantuntijajärjestelmäkehittimen valinta gammalähteiden tunnistamiseen.* Raportti TKK-F-C97, Otaniemi 1988 (in Finnish).
Title in English: *Selecting an Expert System Development Tool for Radionuclide Identification.*

[2] P.A. Aarnio, T.T. Hakulinen, M.P. Huhtinen, *SEITA — Radionuclide Database System.* Report TKK-F-C110, Otaniemi 1989.

[3] P.A. Aarnio, T.T. Hakulinen, “Developing a Personal Computer Based Expert System for Radionuclide Identification”. Article in book *New Computing Techniques in Physics Research.* Proceedings of the First International Workshop on Software Engineering, Artificial Intelligence and Expert Systems in High Energy and Nuclear Physics. March 19–24, 1990, Lyon Villeurbanne, France. Editions du Centre National de la Recherche Scientifique. Paris 1990.

[4] P.A. Aarnio, T.T. Hakulinen, J.T. Routti, “Expert System for Nuclide Identification and Interpretation of Gamma Spectrum Analysis”. Journal of Radioanalytical and Nuclear Chemistry **160** 1 (1992) 245–252.

[5] T.T. Hakulinen, *Knowledge Based System for Identification of Gamma Sources.* Report TKK-F-B147, Otaniemi 1993.

[6] P.A. Aarnio, T.T. Hakulinen, M.P. Huhtinen, “Gamma-Spectrometry and Residual Nuclide Analysis of Pion Irradiated $PbWO_4$ Crystals”. Journal of Radioanalytical and Nuclear Chemistry **248** 2 (2001) 385–393.

[7] S. Arvela, *Update of the Nuclide and Gamma Library of Expert System* SHAMAN. Special assignment in nuclear engineering. Helsinki University of Technology, Department of Engineering Physics and Mathematics 2002.

Selected SAMPO Publications

[8] J.T. Routti, *SAMPO, A Fortran IV Program for Computer Analysis of Gamma Spectra from Ge(Li) Detectors, and for Other Spectra with Peaks.* University of California, Lawrence Berkeley Laboratory, Report UCRL-19452, 1969.

[9] J.T. Routti, S.G. Prussin, “Photopeak Method for the Computer Analysis of Gamma-Ray Spectra from Semiconductor Detectors”. Nuclear Instruments and Methods **72** (1969) 125–142.

[10] G.C. Christensen, M.J. Koskelo, J.T. Routti, *Gamma Spectrum Storage and Analysis Program SAMPO76 with Nuclide Identification.* CERN, Health and Safety Division, Report HS-RP/015/Rev. 1977, Geneva, revised 1978.

[11] M.J. Koskelo, P.A. Aarnio, J.T. Routti, “SAMPO80: An Accurate Gamma Spectrum Analysis Method for Minicomputers”. Nuclear Instruments and Methods in Physics Research **190** (1981) 89–99.

[12] M.J. Koskelo, P.A. Aarnio, J.T. Routti, “SAMPO80: Minicomputer Program for Gamma Spectrum Analysis with Nuclide Identification”. Computer Physics Communications **24** 1 (1981) 11–35.

[13] P.A. Aarnio, J.T. Routti, J.V. Sandberg, “MicroSAMPO — Personal Computer Based Advanced Gamma Spectrum Analysis System”. Journal of Radioanalytical and Nuclear Chemistry **124** 2 (1988) 457–466.

[14] P.A. Aarnio, M.T. Nikkinen, J.T. Routti, "SAMPO 90 — High Resolution Interactive Gamma Spectrum Analysis Including Automation with Macros". Journal of Radioanalytical and Nuclear Chemistry **160** 1 (1992) 289–296.

[15] P.A. Aarnio, M.T. Nikkinen, J.T. Routti, "Gamma Spectrum Analysis Including NAA with SAMPO for Windows". Journal of Radioanalytical and Nuclear Chemistry **193** 1 (1995) 179–185.

[16] P.A. Aarnio, M.T. Nikkinen, J.T. Routti, "UniSampo, Comprehensive Software for Gamma-Spectrum Processing". Journal of Radioanalytical and Nuclear Chemistry **248** 2 (2001) 371–375.

[17] R. Hulmi, *Gammaspektrien synteesi simuloiduilla Ge(Li)-detektorin vastefunktioilla.* Master's thesis. Helsinki University of Technology, Department of Technical Physics 1979 (in Finnish).
Title in English: *Synthesis of Gamma-Ray Spectra Using Simulated Response Functions of Ge(Li) Detectors.*

[18] T. Tynkkynen, *Gammaspektrien syntetisoinnin viimeistelyä.* Special assignment in nuclear engineering. Helsinki University of Technology, Department of Technical Physics 1981 (in Finnish).
Title in English: *Revision of a Gamma-Ray Spectrum Synthesizer.*

[19] J. Laurimaa, *Gammaspektrin syntetisointiohjelma mikrotietokoneympäristöön.* Special assignment in nuclear engineering. Helsinki University of Technology, Department of Information Technology 1990 (in Finnish).
Title in English: *Implementation of a Gamma-Ray Spectrum Synthesizer in a Micro Computer Environment.*

Other Publications

[20] M. Eisenbud, T. Gesell, *Environmental Radioactivity From Natural, Industrial, and Military Sources, Fourth Edition.* Academic Press. San Diego 1997.

[21] Comprehensive Nuclear-Test-Ban Treaty, United Nations General Assembly resolution 50/245. September 10, 1996.

[22] G.F. Knoll, *Radiation Detection and Measurement, Second Edition.* John Wiley & Sons, Inc. New York 1989.

[23] B.R. Kinsey, et al., "The NUDAT/PCNUDAT Program for Nuclear Data". Paper submitted to the 9th International Symposium of Capture Gamma-Ray Spectroscopy and Related Topics, Budapest, Hungary, October 1996. *Data extracted from the PC-NUDAT database version 2.80, August 2002.*
`http://www.nndc.bnl.gov/nndc/nudat/` [cited Feb 6, 2008]

[24] Comprehensive Nuclear-Test-Ban Treaty Organization, `http://www.ctbto.org` [cited Feb 6, 2008]

[25] W.C. Evans, *Model for Assessment of Surveillance Strategies.* PSR Technical Note 1079, Pacific-Sierra Research Corporation 1995.

[26] Provisional Technical Secretariat for the Comprehensive Nuclear-Test-Ban Treaty Organization, *Operational Manual for Radionuclide Monitoring and the International Exchange of Radionuclide Data, Draft.* CTBT/WGB/TL-11/5/Rev.10, Vienna, 7 September 2004.

[27] Senya Ltd., Air Samplers for Radiation Detection, `http://www.senya.fi` [cited Feb 6, 2008]

[28] P. Denier, H. Toivonen, "Evaluation of the International Monitoring System and International Data Centre of the Comprehensive Nuclear-Test-Ban Treaty Organization". Kerntechnik **66** 3 (2001) 147–151.

[29] "Verification Science: The CTBT Noble Gas Verification Component". CTBTO Spectrum 8, July 2006. Available at `http://www.ctbto.org` [cited Feb 6, 2008]

[30] MySQL Open Source Database, `http://www.mysql.com` [cited Feb 6, 2008]

[31] PostgreSQL Open Source Database, `http://www.postgresql.org` [cited Feb 6, 2008]

[32] J. Verplancke, "Low Level Gamma Spectrometry: Low, Lower, Lowest". Nuclear Instruments and Methods in Physics Research **A312** (1992) 174–182.

[33] R. Wordel, D. Mouchel, T. Altzitzoglou, G. Heusser, B.Q. Arnes, P. Meynendonckx, "Study of Neutron and Muon Background in Low-Level Germanium Gamma-Ray Spectrometry". Nuclear Instruments and Methods in Physics Research **A369** (1996) 557–562.

[34] G. Heusser, "Cosmic Ray Interaction Study with Low-Level Ge-Spectrometry". Nuclear Instruments and Methods in Physics Research **A369** (1996) 539–543.

[35] F. Cannizzaro, G. Greco, M. Raneli, M.C. Spitale, E. Tomarchio, "Study of Background Characteristics of a Low-Level HPGe Spectrometer with Passive Shielding in Various Configurations". Nuclear Instruments and Methods in Physics Research **A390** (1997) 167–174.

[36] H. Zhou, X. Wang, C. Wang, M. Hua, G. Huang, G. Fan, T. Lu, "Discrete Gamma Radiation in Interaction of 14.9-MeV Neutrons with Natural Lead". Nuclear Science and Engineering **134** (2000) 106–113.

[37] T. Siiskonen, H. Toivonen, "A Model for Fitting Peaks Induced by Fast Neutrons in an HPGe Detector". Nuclear Instruments and Methods in Physics Research **A540** (2005) 403–411.

[38] E. Bertel, "Beneficial Uses and Production of Isotopes: Status and Trends". Nuclear News **42** 9 (1999) 30–32.

[39] E.J. Peterson, "Accelerator Radioisotopes Save Lives — The Isotope Production Facility at Los Alamos". Los Alamos Science **30** (2006) 112–123.

[40] S.B. Ludwig, J.P. Renier, *Standard- and Extended-Burnup PWR and BWR Reactor Models for the ORIGEN2 Computer Code.* ORNL/TM-11018, Oak Ridge December 1989.

[41] ORIGEN-2.2, Isotope Generation and Depletion Code Matrix Exponential Method, Oak Ridge National Laboratory, available at `http://www.nea.fr/abs/html/ccc-0371.html`. [cited Feb 6, 2008]

[42] T.R. England and B.F. Rider, *Evaluation and Compilation of Fission Product Yields 1993.* Los Alamos National Laboratory Report LA-UR-94-3106, ENDF-349, Oct. 1994.

[43] S. Glasstone, P.J. Dolan, *The Effects of Nuclear Weapons, Third Edition.* The United States Department of Defense and the United States Department of Energy. 1977.

[44] International Atomic Energy Agency, Report by an International Advisory Committee, *The Radiological Situation at the Atolls of Mururoa and Fangataufa — Main Report.* Radiological Assessment Reports Series STI/PUB/1028, Vienna 1998.

[45] T. Siiskonen, H. Toivonen, "Electron Conversion Decay of ^{133m}Xe". Radiation Physics and Chemistry **69** (2004) 23–24.

[46] E. Browne, R.B. Firestone, *Table of Radioactive Isotopes.* John Wiley & Sons, Inc. New York 1986.

[47] L-E De Geer, *CTBT Relevant Radionuclides.* Technical Report PTS/IDC–1999/02, Vienna 1999.

[48] P.W. Krey, "Remote Plutonium Contamination and Total Inventories from Rocky Flats". Health Physics **30** (1976) 209–214.

[49] K. Debertin, R.G. Helmer, *Gamma- and X-Ray Spectrometry with Semiconductor Detectors.* North-Holland. Amsterdam 1988.

[50] G. Gilmore, J. Hemingway, *Practical Gamma-Ray Spectrometry.* John Wiley & Sons, Inc. England 1995.

[51] eV Products, Manufacturer of Cadmium Zinc Telluride detector technology, `http://www.evproducts.com` [cited Feb 6, 2008]

[52] V. Barnes, "GASPAN — An Advanced Computer Code for the Analysis of High Resolution Gamma-Ray Spectra". IEEE Transactions on Nuclear Science **NS-15** 3 (1968) 437–454.

[53] F. Adams, R. Dams, "A Computer Assisted Qualitative Analysis of γ-Ray Spectra". Radiochemical and Radioanalytical Letters **1** 3 (1969) 163–175.

[54] R. Dams, F. Adams, "Automated Identification of Isotopes in Neutron Activated Samples". Journal of Radioanalytical Chemistry **4** (1970) 311–323.

[55] F. Adams, R. Dams, "Computer-Assisted Qualitative Analysis of Gamma-Ray Spectra". Journal of Radioanalytical Chemistry **7** (1971) 329–340.

[56] R. Gunnink, J.B. Niday, *Computerized Quantitative Analysis by Gamma-Ray Spectrometry. Vol. I. Description of the GAMANAL Program.* Report UCRL-51061, Livermore, California 1972.

[57] L.T. Felawka, J.G. Molnar, J.D. Chen, D.G. Boase, *GAMAN — A Computer Program for the Qualitative and Quantitative Evaluation of Ge(Li) Gamma-Ray Spectra.* Report AECL-4217, Pinawa, Manitoba 1973.

[58] H. Weigel, J. Dauk, "A Computer Program for Automatic Gamma-Ray Spectra Analysis with Isotope Identification for the Purpose of Activation Analysis". Journal of Radioanalytical Chemistry **23** (1974) 171–187.

[59] G.W. Phillips, K.W. Marlow, "Automatic Analysis of Gamma-Ray Spectra from Germanium Detectors". Nuclear Instruments and Methods **137** (1976) 525–536.

[60] W. Carder, T.D. Mac Mahon, A. Egan, "IDENT — A Radioisotope-Identification Subroutine for Use with the Gamma-Ray Spectrum Analysis Program SAMPO". Talanta **25** (1978) 21–40.

[61] R. Gunnink, "Computer Techniques for Analysis of Gamma-Ray Spectra". Article in book *Computers in Activation Analysis and Gamma-Ray Spectroscopy.* Proceedings of the ANS Topical Conference at Mayaguez, Puerto Rico, April 30–May 4, 1978. DOE Symposium Series 49, CONF-780421.

[62] G.W. Phillips, "Automatic Computer Analysis of Gamma-Ray Spectra". Article in book *Computers in Activation Analysis and Gamma-Ray Spectroscopy.* Proceedings of the ANS Topical Conference at Mayaguez, Puerto Rico, April 30–May 4, 1978. DOE Symposium Series 49, CONF-780421.

[63] H.P. Yule, "A Computer Code for Qualitative Analysis of Gamma-Ray Spectra". Article in book *Computers in Activation Analysis and Gamma-Ray Spectroscopy.* Proceedings of the ANS Topical Conference at Mayaguez, Puerto Rico, April 30–May 4, 1978. DOE Symposium Series 49, CONF-780421.

[64] L.V. East, E.G. Fisher, S.D. Piner, G.E. Sengstock, "SPECTRAN: A Set of Computer Programs for Analysis of Gamma-Ray Spectra from Ge(Li) Spectrometers". Article in book *Computers in Activation Analysis and Gamma-Ray Spectroscopy.* Proceedings of the ANS Topical Conference at Mayaguez, Puerto Rico, April 30–May 4, 1978. DOE Symposium Series 49, CONF-780421.

[65] T.D. MacMahon, "IDENT-II: An Extension of the Radioisotope Identification Subroutine for Use with the Spectrum Analysis Program SAMPO". Article in book *Computers in Activation Analysis and Gamma-Ray Spectroscopy.* Proceedings of the ANS Topical Conference at Mayaguez, Puerto Rico, April 30–May 4, 1978. DOE Symposium Series 49, CONF-780421.

[66] G.C. Sanderson, "A Least-Squares Method for Computing Nuclide Concentrations from Environmental Gamma-Ray Spectra". Article in book *Computers in Activation Analysis and Gamma-Ray Spectroscopy.* Proceedings of the ANS Topical Conference at Mayaguez, Puerto Rico, April 30–May 4, 1978. DOE Symposium Series 49, CONF-780421.

[67] F. Schley, “CAELCO — An Off-Line Program for Nuclide Identification and Calculation of Element Concentrations from Gamma-Ray Spectra”. Nuclear Instruments and Methods in Physics Research **193** (1982) 141–145.

[68] V.B. Zlokazov, “ACTIV — A Program for Automatic Processing of Gamma-Ray Spectra”. Computer Physics Communications **28** 1 (1982) 27–40.

[69] R.G. Helmer, C.M. McCullagh, “GAUSS VII, a Computer Program for the Analysis of γ-Ray Spectra from Ge Semiconductor Detectors”. Nuclear Instruments and Methods in Physics Research **206** (1983) 477–488.

[70] R.G. Helmer, M.H. Putnam, C.M. McCullagh, “Nuclide Activities Determined from γ-Ray Spectra from Ge-Detectors: A Review with GAUSS VIII as the Example”. Nuclear Instruments and Methods in Physics Research **A242** (1986) 427–436.

[71] W.K. Hensley, E.A. Lepel, M.E. Yuly, K.H. Abel, “Adaptation and Implementation of the RAYGUN Gamma-Ray Analysis Code on the IBM PC”. Journal of Radioanalytical and Nuclear Chemistry **124** 2 (1988) 481–499.

[72] N.B. Kim, Y.S. Kim, J.K. Kim, K.S. Park, “PIAD — A Computer Code for Peak Identification and Activity Determination”. Journal of Radioanalytical and Nuclear Chemistry **125** 2 (1988) 403–417.

[73] R.B. Welch, F. Gyger, D.T. Jost, H.R. von Gunten, U. Krähenbühl, “NEWFIT, a Computer Program for the Analysis of Alpha, X-Ray and Gamma-Ray Spectra”. Nuclear Instruments and Methods in Physics Research **A269** (1988) 615–622.

[74] M. Heimlich, P.A. Beeley, J.A. Page, “GAMANAL-PC: A Program for Gamma-Ray Spectrum Analysis Using a Microcomputer”. Journal of Radioanalytical and Nuclear Chemistry **132** 2 (1989) 281–291.

[75] V. Hnatowicz, V. Ilyushchenko, P. Kozma, “GSAP: Fortran Code for Gamma-Spectrum Analysis”. Computer Physics Communications **60** 1 (1990) 111–125.

[76] R.M. Keyser, “Using Standard Spectra to Develop and Test Gamma-Ray Analysis Software”. Nuclear Instruments and Methods in Physics Research **A286** (1990) 409–414.

[77] A. Baeza, G. Corvo, M. del Rio, C. Miro, J.M. Paniagua, “The Program ESPEC for the Analysis of Gamma Spectra of Environmental Samples”. Applied Radiation and Isotopes **43** 7 (1992) 833–839.

[78] V.S. Kondrashov, Z.D. Moroz, A.A. Kolyshkin, R. Vaillancourt, “A Computer Program for the Identification of Nuclides by Using Median Estimates of Peak Areas in Gamma-Ray Spectra”. Nuclear Instruments and Methods in Physics Research **A328** (1993) 542–546.

[79] M. Blaauw, “The Holistic Analysis of Gamma-Ray Spectra in Instrumental Neutron Activation Analysis”. Nuclear Instruments and Methods in Physics Research **A353** (1994) 269–271.

[80] K. Aleklett, J.-O. Liljenzin, W. Loveland, “DECHAOS — A Program for Automatic or Interactive Analysis of Gamma-Ray Spectra”. Journal of Radioanalytical and Nuclear Chemistry **193** 1 (1995) 187–194.

[81] V.P. Kolotov, V.V. Atrashkevich, “Software ASPRO-NUC: Gamma-Ray Spectrometry, Routine NAA, Isotope Identification and Data Management”. Journal of Radioanalytical and Nuclear Chemistry **193** 2 (1995) 195–206.

[82] K.A. Jayanthi, R.P. Corcuera, G.M.B. Oliveira, *Prototype of an Expert System to Help Nuclide Identification in Gamma Spectrum Analysis.* Report IEAv–032/95, Instituto de Estudos Avançados, Centro Técnico Aeroespacial, São José dos Campos, Brazil, 1995.

[83] K.A. Jayanthi, R.P. Corcuera, G.M.B. Oliveira, *“NIDGAMS”: an Expert System for Nuclide Identification in Gamma-Ray Spectroscopy.* Report IEAv–014/96, Instituto de Estudos Avançados, Centro Técnico Aeroespacial, São José dos Campos, Brazil, 1995.

[84] B. Fazekas, G. Molnár, T. Belgya, L. Dabolczi, A. Simonits, “Introducing HYPERMET-PC for Automatic Analysis of Complex Gamma-Ray Spectra”. Journal of Radioanalytical and Nuclear Chemistry **215** 2 (1997) 271–277.

[85] M.A. Mariscotti, “A Method for Automatic Identification of Peaks in the Presence of Background and its Application to Spectrum Analysis”. Nuclear Instruments and Methods **50** (1967) 309–320.

[86] W. Westmeier, “Background Subtraction in Ge(Li) Gamma-Ray Spectra”. Nuclear Instruments and Methods **180** (1981) 205–210.

[87] G.P. Lasche, R.L. Coldwell, J.A. Nobel, “Detection of Low Levels of Plutonium in Natural Environments from Gamma-Ray Spectra with Advanced Methods in Robust Fitting”. Nuclear Instruments and Methods in Physics Research **A422** (1999) 455–462.

[88] G. Erdtmann, W. Soyka, *The Gamma Rays of the Radionuclides.* Verlag Chemie. New York 1979.

[89] G. Erdtmann, W. Soyka, *The Radionuclide Gamma Ray Data File ”GAMDAT-78”.* Central Division of Chemical Analysis Nuclear Research Establishment, Kernforschungsanlage Jülich GmbH. 1979.

[90] T. Ansaranta, *Parameter Optimization for Automated Gamma Spectrum Analysis.* Special assignment in nuclear engineering. Helsinki University of Technology, Department of Engineering Physics and Mathematics 2000.

[91] A. Leppänen, private communications 1994–1995.

[92] R. Griffiths, “Calculated Total Efficiencies of Coaxial Ge(Li) Detectors”. Nuclear Instruments and Methods **91** (1971) 377–379.

[93] J. Kemppinen, *Germanium-ilmaisimen vasteen simulointi Monte Carlo -menetelmällä.* Special assignment in nuclear engineering. Helsinki University of Technology, Department of Engineering Physics and Mathematics 2003.
Title in English: *Monte Carlo Simulation of the Response of a Germanium Detector.*

[94] L.A. Currie, "Limits for Qualitative Detection and Quantitative Determination — Application to Radiochemistry". Analytical Chemistry **40** 3 (1968) 586–593.

[95] I.V. Aničin, C.T. Yap, "New Approach to Detection Limit Determination in Spectroscopy". Nuclear Instruments and Methods in Physics Research **A259** (1987) 525–528.

[96] *Nachweisgrenze und Erkennungsgrenze bei Kernstrahlungsmessungen; Zählende spektrometrische Messungen ohne Berücksichtigung des Probenbehandlungseinflusses.* Normenausschuß Kerntechnik (NKe) im DIN Deutsches Institut für Normung e.V., Deutsche Norm DIN 25 482 Teil 2, 1992.
Title in English: *Limit of Detection and Limit of Decision for Nuclear Radiation Measurements; Counting Spectrometric Measurements, Neglecting the Influence of Sample Treatment.*

[97] *Nachweisgrenze und Erkennungsgrenze bei Kernstrahlungsmessungen; Zählende hochauflösende gammaspektrometrische Messungen ohne Berücksichtigung des Probenbehandlungseinflusses.* Normenausschuß Kerntechnik (NKe) im DIN Deutsches Institut für Normung e.V., Deutsche Norm DIN 25 482 Teil 5, 1993.
Title in English: *Limit of Detection and Limit of Decision for Nuclear Radiation Measurements; Counting Measurements by High Resolution Gamma Spectrometry, Neglecting the Influence of Sample Treatment.*

[98] K. Heydorn, S.K. Rhee, "Advance Prediction of Single- and Double-Escape Peak Areas in Gamma-Ray Spectrometry". ANS Transactions **60** (1989) 3–4.

[99] L.C. Maximon, "Simple Analytic Expressions for the Total Born Approximation Cross Section for Pair Production in a Coulomb Field". Journal of Research of the National Bureau of Standards **72B** 1 (1968) 79–88.

[100] J.H. Chao, "Single- to Double-Escape Peak Ratio of Germanium Detectors". Nuclear Instruments and Methods in Physics Research **A317** (1992) 537–540.

[101] D.S. Andreev, K.I. Erokhina, V.S. Zvonov, I.Kh. Lemberg, "Consideration of Cascade Transitions in Determining the Absolute Yield of Gamma Rays". Instruments and Experimental Techniques **25** (1972) 1358–1360.

[102] G.J. McCallum, G.E. Coote, "Influence of Source-Detector Distance on Relative Intensity and Angular Correlation Measurements with Ge(Li) Spectrometers". Nuclear Instruments and Methods **130** (1975) 189–197.

[103] K. Debertin, U. Schötzig, "Coincidence Summing Corrections in Ge(Li)-Spectrometry at Low Source-to-Detector Distances". Nuclear Instruments and Methods **158** (1979) 471–477.

[104] K. Sinkko, H. Aaltonen, *Calculation of the True Coincidence Summing Correction for Different Sample Geometries in Gamma-Ray Spectroscopy.* Report STUK-B-VALO 40, Helsinki 1985.

[105] R.J. Gehrke, R.G. Helmer, R.C. Greenwood, “Precise Relative Gamma Ray Intensities for Calibration of Ge Semiconductor Detectors”. Nuclear Instruments and Methods **147** (1977) 405–423.

[106] F.J. Schima, D.D. Hoppes, “Tables for Cascade-Summing Corrections in Gamma-Ray Spectrometry”. International Journal of Applied Radiation and Isotopes **34** 8 (1983) 1109–1114.

[107] J. Morel, B. Chauvenet, A. Kadachi, “Coincidence-Summing Corrections in Gamma-Ray Spectrometry for Normalized Geometries”. International Journal of Applied Radiation and Isotopes **34** 8 (1983) 1115–1122.

[108] T.M. Semkow, G. Mehmood, P.P. Parekh, M. Virgil, “Coincidence Summing in Gamma-Ray Spectroscopy”. Nuclear Instruments and Methods in Physics Research **A290** (1990) 437–444.

[109] M. Korun, R. Martinčič, “Coincidence Summing in Gamma and X-Ray Spectrometry”. Nuclear Instruments and Methods in Physics Research **A325** (1993) 478–484.

[110] M. Korun, R. Martinčič, “Activity Calculation for Voluminous Samples in the Presence of Coincidence Summing Effects”. Nuclear Instruments and Methods in Physics Research **A355** (1995) 600–608.

[111] L-E De Geer, *X-X and Other True Coincidence Summation Peaks in Air Filter Gamma Spectra Dominated by the Lead-212 Decay Chain.*
Report CTBT/PTS/TP/2003-1, Vienna 2003.

[112] C.M. Lederer, V.S. Shirley (eds.), *Table of Isotopes, Seventh Edition.* John Wiley & Sons, Inc. New York 1978.

[113] M.R. Bhat, “Evaluated Nuclear Structure Data File (ENSDF)”. *Nuclear Data for Science and Technology.*, page 817, ed. S.M. Qaim, Springer Verlag, Berlin, Germany, 1992. *Data extracted from the ENSDF database at the National Nuclear Data Center in 1992.*

[114] R.B Firestone, *Table of Isotopes, Eighth Edition.* John Wiley & Sons, Inc. New York 1996.

[115] R.B Firestone, *Table of Isotopes, Eighth Edition, 1999 Update.* John Wiley & Sons, Inc. New York 1999.

[116] K. Saloriutta, *Hajoamiskaavioiden rakentaminen ENSDF-tietokannasta.* Special assignment in nuclear engineering. Helsinki University of Technology, Department of Technical Physics and Mathematics 2004.
Title in English: *Building Decay Schemes from ENSDF Database.*

[117] M.P. Fioratti, S.R. Piermattei, “Calculation of the Escape Peak for Ge-Li and NaI Radiation Detectors”. Nuclear Instruments and Methods **96** (1971) 605–608.

[118] S.J.B. Reed, N.G. Ware, "Escape Peaks and Internal Fluorescence in X Ray Spectra Recorded with Lithium Drifted Silicon Detectors". Journal of Physics E **5** (1972) 582–584.

[119] J.S. Hansen, J.C. McGeorge, D. Nix, W.D. Schmidt-Ott, I. Unus, R.W. Fink, "Accurate Efficiency Calibration and Properties of Semiconductor Detectors for Low-Energy Photons". Nuclear Instruments and Methods **106** (1973) 365–379.

[120] S.M. Seltzer, "Calculated Response of Intrinsic Germanium Detectors to Narrow Beams of Photons with Energies up to ∼300 keV". Nuclear Instruments and Methods in Physics Research **188** (1981) 133–151.

[121] K. Ilakovac, V. Horvat, N. Ilakovac, "The Dependence of the Escape of Characteristic X-Rays from Planar Germanium Detectors on the Direction of the Incident Radiation". Nuclear Instruments and Methods in Physics Research **228** (1984) 210–214.

[122] M.J. Berger, J.H. Hubbell, *XCOM: Photon Cross Sections Database.* NIST Standard Reference Database 8 (XGAM), NBSIR 87-3597.
`http://physics.nist.gov/PhysRefData/Xcom/Text/XCOM.html` [cited Feb 6, 2008]

[123] H. Toivonen, M. Nikkinen, private communications, October 1998–February 1999.

[124] MCNP — A General Monte Carlo N-Particle Transport Code. Los Alamos National Laboratory. `http://mcnp-green.lanl.gov` [cited Feb 6, 2008]

[125] H. Bateman, "Solution of a System of Differential Equations Occurring in the Theory of Radioactive Transformations". Proceedings of Cambridge Philosophical Society, **15** (1910) 423–427.

[126] F. Hayes-Roth, D.A. Waterman, D.B. Lenart (eds.), *Building Expert Systems.* Addison-Wesley Publishing Company, Inc. 1983.

[127] J.L. Alty, M.J. Coombs, *Expert Systems — Concepts and Examples.* NCC Publications. Manchester 1984.

[128] P. Harmon, D. King, *Expert Systems.* John Wiley & Sons, Inc.. New York 1985.

[129] *GoldWorks 286 Version 1.1 Documentation.* Gold Hill Computers. 1988.

[130] *American National Standard for Information Systems — Programming Language C.* Document ANSI X3J11/88–159, Washington 1989.

[131] B.R. Kinsey, et al., "The NUDAT/PCNUDAT Program for Nuclear Data". Paper submitted to the 9th International Symposium of Capture Gamma-Ray Spectroscopy and Related Topics, Budapest, Hungary, October 1996. *Data extracted from the NUDAT database at the National Nuclear Data Center in 1992.*

[132] S.Y.F. Chu, L.P. Ekström and R.B. Firestone, *WWW Table of Radioactive Isotopes, database version 2/28/1999.*
`http://nucleardata.nuclear.lu.se/nucleardata/toi/` [cited Feb 6, 2008]

[133] William C. Evans, private communications, October 1996 – December 1997.

[134] T.W. Burrows, private communications 2002.

[135] *Preliminary User's Guide for Radionuclide Data Processing Software, Version 1.2.* Pacific-Sierra Research Corporation, Arlington 1996.

[136] L.R. Mason, D.L. Williams, W.C. Evans, J.D. Bohner, J.B. Bohlin, K.K. Keksz, E.A. Shnekendorf, *Radionuclide Monitoring Operations Report of the Prototype International Data Center — Third Quarter CY1996.* Pacific-Sierra Research Corporation Technical Report 2682, Arlington 1996.

[137] L.R. Mason, J.D. Bohner, W.C. Evans, J.B. Bohlin, D.L. Williams, K.K. Keksz, *Radionuclide Monitoring Operations Report of the Prototype International Data Center — Fourth Quarter CY1996.* Pacific-Sierra Research Corporation Technical Report 2698, Arlington 1997.

[138] L.R. Mason, J.D. Bohner, W.C. Evans, *Radionuclide Monitoring Operations Report of the Prototype International Data Center — First Quarter CY1997.* Pacific-Sierra Research Corporation Technical Report 2715, Arlington 1997.

[139] J.D. Bohner, L.R. Mason, S.R. Biegalski, *Radionuclide Monitoring Operations Report of the Prototype International Data Center — Second Quarter CY1997.* Pacific-Sierra Research Corporation Technical Report 2726, Arlington 1997.

[140] John D. Bohner, private communications, September 1997.

[141] J. Kang, F.N. von Hippel, "U-232 and the Proliferation-Resistance of U-233 in Spent Fuel". Science & Global Security, **9** (2001) 1–32.

[142] H. Toivonen, S. Klemola, J. Lahtinen, A. Leppänen, R. Pöllänen, A. Kansanaho, A.L. Savolainen, A.Sarkanen, I. Valkama, M. Jäntti, *Radioactive Release from Sosnovyy Bor, St. Petersburg, in March 1992.* Report STUK-A104, Helsinki 1992.

[143] H. Toivonen, R. Pöllänen, A. Leppänen, S. Klemola, J. Lahtinen, K. Servomaa, A.L. Savolainen, I. Valkama, "A Nuclear Incident at a Power Plant in Sosnovyy Bor, Russia". Health Physics **63** 5 (1992) 571–573.

[144] H. Toivonen, R. Pöllänen, A. Leppänen, S. Klemola, J. Lahtinen, "Release from the Nuclear Power Plant in Sosnovyy Bor in March 1992". Radiochimica Acta **57** (1992) 169–172.

[145] A. Leppänen, *Radionuclides in Ground-Level Air — Quarterly results of air surveillance programme, first quarter 1992.* STUK Aerosol Laboratory Report No 30, Helsinki 1992.

[146] Preparatory Commission for the Comprehensive Nuclear-Test-Ban Treaty Organization, *Recommended Standard List of Relevant Radionuclides for IDC Event Screening.* CTBT/WGB/TL-2/40, Report of Informal Workshop, Melbourne, Australia. January 17–21, 2000.

[147] P. Zagyvai, R.M. Parr, L.G. Nagy, "Additional Results for the "G-1" IAEA Intercomparison of Methods for Processing Ge(Li) Gamma-Ray Spectra". Journal of Radioanalytical and Nuclear Chemistry **89** 2 (1985) 589–607.

[148] C.G. Sanderson, "An Evaluation of Commercial IBM PC Software for the Analysis of Low Level Environmental Gamma-Ray Spectra". Environment International **14** (1988) 379–384.

[149] K. Heydorn, L.H. Christensen, "Verification Testing of Commercially Available Computer Programs for Photopeak Area Evaluation — The Influence of Peak Width". Journal of Radioanalytical and Nuclear Chemistry **124** 2 (1988) 467–480.

[150] M.J. Koskelo, M.T. Mercier, "Verification of Gamma Spectroscopy Programs: a Standardized Approach". Nuclear Instruments and Methods in Physics Research **A299** (1990) 318–321.

[151] K.M. Decker, C.G. Sanderson, "A Reevaluation of Commercial IBM PC Software for the Analysis of Low-Level Environmental Gamma-Ray Spectra". Applied Radiation and Isotopes **43** 1/2 (1992) 323–337.

[152] M. Blaauw, V. Osorio Fernandez, P. van Espen, G. Bernasconi, R. Capote Noy, H. Manh Dung, N.I. Molla, "The 1995 IAEA Intercomparison of γ-Ray Spectrum Analysis Software". Nuclear Instruments and Methods in Physics Research **A387** (1997) 416–432.

[153] *Intercomparison of Gamma Ray Analysis Software Packages.* IAEA TECDOC-1011, Austria 1998.

[154] Mohammad Wasim, private communications, September 1997.

[155] J.S. Milton, J.C. Arnold, *Probability and Statistics in the Engineering and Computing Sciences.* McGraw-Hill Book Co. Singapore 1986.

[156] IDC Documentation, IDC Processing of Radionuclide Data, Revision 2. Report IDC-5.2.2Rev2, December 2001.

[157] Preparatory Commission for the Comprehensive Nuclear-Test-Ban Treaty Organization, *Report of the IDC Technical Experts Meeting on Event Screening.* CTBT/WGB/TL-2/10, Vienna, Austria. August 11, 1998.

[158] E. Arik, *CTBTO Peak Identification List.* CTBTO Draft Report, Vienna, October 2000.

[159] H. Toivonen et al., *Peak Significance for IDC Radionuclide Software.* CTBTO Draft Report, Vienna, February 2000.

Appendix A

SHAMAN's Gamma-Ray and Nuclide Lists for Air Filter Spectra

In the following, lists are presented of gamma-rays and nuclides handled in a special way by expert system SHAMAN (v. 1.16) when analyzing air filter spectra:

- Gamma-rays with a user-given explanation are presented in Table A.1. Peaks within the lookup tolerance of these energies and with a suitable count rate are explained with the explanation given here, i.e., no gamma lines are searched for in the SHAMAN reference library. The gamma-rays fall into two main categories: (a) gamma-rays from isomers and excitation states produced by cosmic radiation in the detector and its shield, and (b) X-ray sum peaks of ^{220}Rn daughters that the current version of SHAMAN cannot calculate.
- The nuclides declared as unacceptable in air filter spectra are presented in Table A.2. These nuclides have gamma lines close to common air filter spectrum peaks and were therefore spuriously identified by SHAMAN in many test runs. These nuclides are unlikely to be seen in air filter samples and they are irrelevant for CTBT monitoring.
- The nuclides declared as CTBT relevant are presented in Table A.3. SHAMAN calculates the MDA (minimum detectable activity) values for these nuclides if they are not identified in an air filter spectrum. The list has been modified from the "official" one [146] by adding a few gamma-emitting daughter nuclides with a shorter or a slightly longer half-life. In SHAMAN's reference library, gamma lines of these daughters are separated from the parent's gamma lines, which may be in contrast to a tailored library utilized by other software packages. Thus, identification of either parent or daughter would be CTBT relevant.

  The CTBT relevance is actually utilized by a SHAMAN postprocessor that categorizes air filter spectra for their importance for Treaty verification purposes, using a scale from 1 to 5 as introduced in Ref. [157]. In essence, spectrum identification results including one relevant nuclide are given category 4 and those including multiple relevant nuclides, of which at least one fission product, category 5.

Gamma-ray energy (keV)	Explanation and references
(53.53)	^{73m}Ge primary photopeak [32, 34]
66.72	^{73m}Ge sum peak [32, 34, 35]
109.89	^{19*}F excitation peak [133]
(139.68)	^{75m}Ge primary photopeak [32–34]
149.66	^{212}Pb X+X sum peak [111, 158]
151.96	^{212}Pb X+X sum peak [111, 158]
154.26	^{212}Pb X+X sum peak [111, 158]
159.7	^{77m}Ge IT-photopeak [32, 33]
162.16	^{212}Pb X+X sum peak [111, 158]
164.46	^{212}Pb X+X sum peak [111, 158]
166.90	^{212}Pb X+X sum peak [111]
190.00	^{212}Pb X-ray sum peak [111, 158]
192.29	^{212}Pb X-ray sum peak [111, 158]
198.39	^{71m}Ge sum peak [32–35]
202.36	^{212}Pb X-ray sum peak [111]
374.90	^{212}Pb X-ray sum peak [111, 133]
377.20	^{212}Pb X-ray sum peak [111, 133]
387.27	^{212}Pb X-ray sum peak [111, 133]
389.88	^{212}Pb X-ray sum peak [111, 133]
537.45	^{206*}Pb excitation peak [36, 158]
558.45	^{114*}Cd excitation peak [32, 158]
(569.70)	^{207m}Pb primary photopeak [32, 35, 36]
595.85	^{74*}Ge excitation peak [32–34, 159]
655.99	^{208}Tl X-ray sum peak [111, 133]
658.16	^{208}Tl X-ray sum peak [111, 133]
667.97	^{208}Tl X-ray sum peak [111, 133]
669.62	^{63*}Cu excitation peak [32–35, 158]
670.50	^{208}Tl X-ray sum peak [111, 133]
691.55	^{72*}Ge excitation peak [32–34, 159]
794.59	^{208}Tl X-ray sum peak [111]
803.10	^{206*}Pb excitation peak [32, 35, 36, 158]
816.55	^{208}Tl X-ray sum peak [111]
834.14	^{72*}Ge excitation peak [34, 158]
835.80	^{208}Tl sum peak [111, 133]
843.76	^{27*}Al excitation peak
846.77	^{56*}Fe excitation peak [34, 35, 158, 159]
881.01	^{206*}Pb excitation peak [36]
899.15	^{204*}Pb excitation peak
962.06	^{63*}Cu excitation peak [32, 34, 35, 158]
1014.45	^{27*}Al excitation peak
1022.00	Double annihilation peak [158, 159]
(1063.66)	^{207m}Pb primary photopeak [36]
1115.55	^{65*}Cu excitation peak [32, 35, 158, 159]
1869.88	^{208}Tl sum peak's DE peak [111]
2175.72	^{208}Tl sum peak's DE peak [111]
2223.245	^{2}H excitation peak [34, 158, 159]
2380.88	^{208}Tl sum peak's SE peak [111]
2453.10	^{208}Tl sum peak's DE peak [111]
2686.72	^{208}Tl sum peak's SE peak [111, 133]
2687.34	^{208}Tl X-ray sum peak [111, 133]
2689.50	^{208}Tl X-ray sum peak [111, 133]
2699.31	^{208}Tl X-ray sum peak [111, 133]
2701.88	^{208}Tl X-ray sum peak [111]
2891.89	^{208}Tl sum peak [133]
3125.30	^{208}Tl sum peak [111, 133, 158]

Table A.1: Air filter gamma-rays with a user-given explanation. The energies in parentheses are explained by SHAMAN when it identifies the corresponding nuclide.

Nuclide	Primary gamma-ray (keV)	Nuclide	Primary gamma-ray (keV)
^{40}Cl	1461.0	^{163m}Ho	299.0
^{61}Co	67.415	^{166}Ho	80.574
^{57}Ni	1377.63	^{172}Er	610.062
^{64}Cu	1345.77	^{170}Tm	84.255
^{67}Cu	184.577	^{172}Tm	1093.59
^{71}Zn	910.3	^{174m}Lu	67.058
^{72}Zn	144.7	^{178}Lu	93.179
^{68}Ga	1077.34	^{173}Hf	123.675
^{78}Ge	277.3	^{177n}Hf	277.3
^{73}As	53.437	^{179n}Hf	453.59
^{77}As	239.011	^{183}Hf	73.173
^{78}As	613.8	^{180}Ta	93.4
^{73}Se	67.07	^{182}Ta	67.75
^{80m}Br	37.052	^{184}Ta	414.01
^{77}Kr	129.64	^{185}W	125.358
^{79m}Kr	130.01	^{188}W	63.582
^{83}Sr	762.65	^{183}Re	162.327
^{95m}Nb	235.69	^{184m}Re	252.845
^{98m}Nb	722.626	^{190}Re	186.68
^{95}Tc	765.789	^{190m}Re	186.68
^{103m}Rh	39.755	^{183}Os	381.76
^{105m}Rh	129.57	^{191}Os	129.431
^{106m}Rh	450.8	^{191m}Os	74.38
^{101}Pd	296.29	^{192m}Os	205.794
^{103}Pd	39.748	^{193}Os	138.92
^{111}Pd	70.43	^{191m}Ir	129.431
^{107m}Ag	93.124	^{193m}Ir	80.22
^{109m}Ag	88.034	^{194}Ir	328.448
^{113}Ag	298.6	^{195}Ir	98.85
^{107}Cd	93.124	^{195m}Ir	98.85
^{109}Cd	88.034	^{189}Pt	721.38
^{111m}In	537.0	^{195m}Pt	98.9
^{116m}In	1293.54	^{197}Pt	77.35
^{119}In	763.14	^{195}Au	98.88
^{117m}Sn	158.56	^{196m}Au	84.66
^{126}Sn	87.567	^{199}Au	158.379
^{119m}Te	153.59	^{197}Hg	77.351
^{123m}Te	158.96	^{199m}Hg	158.3
^{133m}Te	912.671	^{208}Bi	2610.0
^{133m}I	912.0	^{211}Bi	351.06
^{120}Xe	72.6	^{215}Bi	293.54
^{127m}Xe	124.6	^{218}Rn	609.31
^{129m}Xe	39.578	^{225}Ra	40.0
^{134m}Cs	127.502	^{227}Th	235.96
^{137m}Ba	661.657	^{231}Th	84.214
^{144m}Pr	59.03	^{228}Pa	911.2
^{149}Nd	211.309	^{232}Pa	969.315
^{145}Pm	72.4	^{239}U	74.664
^{145}Sm	61.22	^{240}U	66.5
^{152n}Eu	89.847	^{243}Pu	84.0
^{158}Tb	98.918	^{245}Pu	327.428
^{160}Tb	298.58	^{242}Am	42.13
^{161}Tb	74.567	^{242m}Am	48.63
^{163}Tb	351.2	^{243}Am	74.66
^{161m}Ho	211.15	^{245}Am	252.72
^{162m}Ho	185.0		

Table A.2: List of unacceptable nuclides in air filter spectra used by Shaman v. 1.16.

Fission products		**Activation products**	
Nuclide	Explanation and references	Nuclide	Explanation and references
^{91}Sr	[47, 146, 157]	^{24}Na	[47, 146]
^{91m}Y	Shorter lived daughter [157]	^{42}K	[47, 146]
^{91}Y	[47, 146]	^{46}Sc	[47, 146]
^{93}Y	[47, 146, 157]	^{47}Sc	[47, 146]
^{95}Zr	[47, 146, 157]	^{51}Cr	[47, 146]
^{95}Nb	[47, 146, 157]	^{54}Mn	[47, 146]
^{97}Zr	[47, 146, 157]	^{57}Co	[47, 146]
^{97}Nb	Shorter lived daughter [157]	^{58}Co	[47, 146]
^{99}Mo	[47, 146, 157]	^{59}Fe	[47, 146]
^{99m}Tc	[47, 146, 157]	^{60}Co	[47, 146]
^{103}Ru	[47, 146, 157]	^{65}Zn	[47, 146]
^{105}Rh	[47, 146, 157]	^{69m}Zn	[47, 146]
^{106}Ru	[47, 146, 157]	^{72}Ga	[47, 146]
^{106}Rh	Shorter lived daughter [157]	^{74}As	[47, 146]
^{111}Ag	[47, 146, 157]	^{76}As	[47, 146]
^{112}Pd	[47, 146, 157]	^{84}Rb	[47, 146]
^{112}Ag	Shorter lived daughter [47, 146, 157]	^{86}Rb	[47, 146]
^{115m}Cd	[47, 146, 157]	^{88}Y	[47, 146]
^{115}Cd	[47, 146, 157]	^{89}Zr	[47, 146]
^{115m}In	Shorter lived daughter [157]	^{102}Rh	[47, 146]
^{125}Sn	[47, 146, 157]	^{106m}Ag	[47, 146]
^{125}Sb	[47, 146, 157]	^{108m}Ag	[47, 146]
^{126}Sb	[47, 146, 157]	^{110m}Ag	[47, 146]
^{127}Sb	[47, 146, 157]	^{120b}Sb	[47, 146]
^{128}Sb	[47, 146, 157]	^{122}Sb	[47, 146]
^{129m}Te	[47, 146, 157]	^{124}Sb	[47, 146]
^{129}Te	Shorter lived daughter [157]	^{132}Cs	[47, 146]
^{130}I	[47, 146]	^{133}Ba	[47, 146]
^{131m}Te	[47, 146, 157]	^{134}Cs	[47, 146, 157]
^{131}Te	Shorter lived daughter [157]	^{152m}Eu	[47, 146]
^{131}I	[47, 146, 157]	^{152}Eu	[47, 146]
^{132}Te	[47, 146, 157]	^{168}Tm	[47, 146]
^{132}I	Shorter lived daughter [47, 146, 157]	^{187}W	[47, 146]
^{133}I	[47, 146, 157]	^{190}Ir	[47, 146]
^{133m}Xe	Longer lived daughter [157]	^{192}Ir	[47, 146]
^{133}Xe	Longer lived daughter [157]	^{196}Au	[47, 146]
^{135}I	[47, 146, 157]	^{196n}Au	[47, 146]
^{135m}Xe	Shorter lived daughter [157]	^{198}Au	[47, 146]
^{135}Xe	Longer lived daughter [157]	^{203}Pb	[47, 146]
^{136}Cs	[47, 146, 157]	^{224}Ra	[47, 146]
^{137}Cs	[47, 146, 157]	^{237}U	[47, 146]
^{140}Ba	[47, 146, 157]	^{239}Np	[47, 146]
^{140}La	[47, 146, 157]	^{241}Am	[47]
^{141}Ce	[47, 146, 157]		
^{143}Ce	[47, 146, 157]		
^{144}Ce	[47, 146, 157]		
^{144}Pr	Shorter lived daughter [157]		
^{147}Nd	[47, 146, 157]		
^{149}Pm	[47, 146, 157]		
^{151}Pm	[47, 146, 157]		
^{153}Sm	[47, 146, 157]		
^{155}Eu	[47, 146, 157]		
^{156}Sm	[47, 146, 157]		
^{156}Eu	[47, 146, 157]		
^{157}Eu	[47, 146, 157]		

Table A.3: Nuclides declared in Shaman as CTBT relevant in air filter spectrum analysis.

Appendix B

Details on the Novaya Zemlya and Sosnovyy Bor Gamma-Ray Spectra

In the following, details on the gamma-ray spectra measured by the Finnish radiation monitoring network after the Novaya Zemlya release in August 1987 and the Sosnovyy Bor release in March 1992 are shown in tabular format. The tables present:

- Basic information on the test set of 34 Novaya Zemlya spectra.
- The anthropogenic nuclides present in 34 Novaya Zemlya spectra. The nuclides flagged with *y* were identified by SHAMAN v. 1.11 in May 2003, those flagged with *n* were not.
- Basic information on the test set of 65 Sosnovyy Bor spectra.
- The anthropogenic nuclides present in 65 Sosnovyy Bor spectra. The nuclides flagged with *y* were identified by SHAMAN v. 1.11 in May 2003, those flagged with *n* were not.

The term "nuclides present" above is to be understood as the nuclide list that is needed for a complete explanation of the peaks that have been found in each spectrum. These identification results are based on a careful interactive analysis of the spectra utilizing software packages SAMPO and SHAMAN. Naturally, if the peak analysis is made with a different search threshold or critical limit, leading to a different peak list, the list of nuclides to be identified will reflect the changes.

For detailed discussion, see Secs. 6.2 and 6.3.

Spectrum identifier	Middle of sampling	Decay time	Live time	Total volume	Station
59987f5	1987/08/06 20:00	5.906 d	38.75 m	3 500 m^3	Ivalo
42387n3	1987/08/08 20:00	8.576 d	23.82 m	3 500 m^3	Ivalo
63887f6	1987/08/08 12:00	3.483 d	21.99 m	8 050 m^3	Rovaniemi
64487f6	1987/08/08 08:00	6.289 d	1.07 h	73 783 m^3	Helsinki
59587f5	1987/08/10 10:45	3.833 h	24.31 m	4 240 m^3	Helsinki
46087n1	1987/08/11 23:25	1.465 d	21.96 m	805 m^3	Helsinki (coal)
59787f5	1987/08/11 00:00	23.170 h	7.33 m	2 070 m^3	Rovaniemi
64187f6	1987/08/11 23:23	1.467 d	23.47 m	52 022 m^3	Helsinki
60587f2	1987/08/12 00:45	1.304 d	24.67 m	1 970 m^3	Rovaniemi
60187f5	1987/08/13 08:18	1.224 d	1.09 h	7 192 m^3	Helsinki
60787f2	1987/08/13 00:40	1.349 d	1.17 h	1 990 m^3	Rovaniemi
64787f6	1987/08/14 00:58	3.308 d	6.52 m	1 930 m^3	Rovaniemi
60587f5	1987/08/15 08:35	2.128 d	20.70 m	72 735 m^3	Helsinki
61187f2	1987/08/15 08:35	2.128 d	21.01 m	1 158 m^3	Helsinki (coal)
60687f5	1987/08/16 00:45	2.323 d	22.69 m	5 930 m^3	Rovaniemi
61287f2	1987/08/16 08:25	2.157 d	18.90 m	15 089 m^3	Helsinki
61887f2	1987/08/18 19:45	1.660 d	19.82 m	853 m^3	Helsinki (coal)
65387f6	1987/08/18 00:36	1.350 d	23.56 m	1 950 m^3	Rovaniemi
65587f6	1987/08/18 19:45	1.763 d	18.81 m	55 443 m^3	Helsinki
61187f5	1987/08/19 00:35	1.326 d	6.27 m	1 950 m^3	Rovaniemi
61987f5	1987/08/19 20:35	1.790 d	1.07 h	10 861 m^3	Helsinki
62187f5	1987/08/20 12:30	3.917 d	20.89 m	3 880 m^3	Rovaniemi
62387f5	1987/08/22 07:45	3.052 d	30.06 m	1 173 m^3	Helsinki (coal)
66087f6	1987/08/22 20:45	1.733 d	17.07 m	10 882 m^3	Helsinki
66287f6	1987/08/22 07:45	3.052 d	23.52 m	73 656 m^3	Helsinki
62887f5	1987/08/23 00:35	5.582 d	1.10 h	5 880 m^3	Rovaniemi
62787f5	1987/08/25 19:50	1.800 d	23.46 m	872 m^3	Helsinki (coal)
66987f6	1987/08/25 20:55	1.756 d	23.53 m	10 849 m^3	Helsinki
67087f6	1987/08/25 19:50	2.783 d	1.10 h	55 332 m^3	Helsinki
64287f2	1987/08/26 00:40	2.582 d	1.10 h	5 860 m^3	Rovaniemi
63587f5	1987/08/29 12:35	2.931 d	25.22 m	7 830 m^3	Rovaniemi
65087f2	1987/08/29 12:35	2.937 d	25.03 m	1 128 m^3	Helsinki (coal)
67687f6	1987/08/29 07:50	3.128 d	25.31 m	73 054 m^3	Helsinki
69087f6	1987/09/01 20:00	2.842 d	1.06 h	55 635 m^3	Nurmijärvi

Table B.1: Basic information on the test set of 34 Novaya Zemlya spectra.

Table B.2: Anthropogenic nuclides present in 34 Novaya Zemlya spectra. The nuclides flagged with *y* were identified by SHAMAN, those flagged with *n* were not.

Spectrum identifier	^{95}Zr	^{95}Nb	^{99m}Tc	^{103}Ru	^{106}Rh	^{122}Sb	^{125}Sb	^{132}Te	^{131}I	^{132}I	^{133}I	^{133}Xe	^{134}Cs	^{137}Cs	^{140}La	^{144}Ce	^{144}Pr	^{224}Ra	Total number
59987f5			*y*	*y*		*y*	*y*	*y*	*y*	*y*	*y*	*y*		*y*					10
42387n3								*y*	*y*	*y*		*y*							4
63887f6			*y*	*y*				*y*	*y*	*y*	*y*	*y*	*y*	*y*	*y*				10
64487f6													*y*	*y*					2
59587f5													*y*	*y*					2
46087n1									*y*										1
59787f5							*y*		*y*					*y*					3
64187f6								*y*	*y*	*y*			*y*	*y*					5
60587f2								*y*	*y*	*n*								*y*	4
60187f5							*y*	*y*	*y*	*y*			*y*	*y*					6
60787f2									*y*										1
64787f6									*y*										1
60587f5								*y*	*y*	*y*			*y*	*y*					5
61187f2									*y*			*n*							2
60687f5								*y*	*y*	*n*				*y*					4
61287f2									*y*				*y*	*y*					3
61887f2									*y*										1
65387f6									*y*				*y*	*y*					3
65587f6									*y*	*y*			*y*	*y*					4
61187f5									*y*										1
61987f5							*n*		*y*				*y*	*y*					4
62187f5							*y*		*y*					*y*					3
62387f5		*y*							*y*			*n*						*y*	4
66087f6													*y*	*y*					2
66287f6									*y*				*y*	*y*					3
62887f5							*y*		*y*				*y*	*y*					4
62787f5							*n*		*y*			*y*		*y*				*y*	5
66987f6									*y*				*y*	*y*					3
67087f6									*y*				*y*	*y*					3
64287f2									*y*				*y*	*y*					3
63587f5							*y*							*y*				*y*	3
65087f2																			0
67687f6	*y*	*y*			*y*				*y*				*y*	*y*		*y*	*y*		8
69087f6									*y*				*y*	*y*		*y*			4
Total	1	2	2	2	1	1	8	8	29	9	2	6	18	24	1	2	1	4	121

Spectrum identifier	Middle of sampling	Decay time	Live time	Total volume	Station
13692f2	1992/02/20 12:00	33.150 d	18.67 h	1 m^3	Helsinki (coal)
11592f7	1992/03/05 21:24	7.758 d	2.71 d	151 307 m^3	Helsinki
07292k5	1992/03/11 08:05	14.370 d	17.90 h	4 000 m^3	Imatra (coal)
15092f7	1992/03/16 06:14	21.420 d	18.09 h	17 749 m^3	Viitasaari (coal)
05392f1	1992/03/17 20:25	26.780 d	17.90 h	7 948 m^3	Nurmijärvi (coal)
09392f8	1992/03/19 18:18	129.700 d	1.98 d	27 114 m^3	Viitasaari
04892f1	1992/03/19 19:43	14.770 d	2.80 d	27 618 m^3	Imatra
07392f1	1992/03/19 20:46	124.800 d	2.00 d	124 219 m^3	Nurmijärvi
06592f8	1992/03/19 20:56	14.720 d	2.80 d	137 801 m^3	Helsinki
15092f6	1992/03/21 07:50	3.118 d	1.01 d	15 612 m^3	Imatra
13592f7	1992/03/23 20:30	1.547 d	9.06 h	4 093 m^3	Imatra
13092f5	1992/03/23 21:14	13.750 h	3.08 h	21 837 m^3	Helsinki
07792n4	1992/03/24 01:10	8.492 d	20.57 h	5 634 m^3	Imatra
13492f7	1992/03/24 11:40	5.317 h	16.64 h	780 m^3	Imatra
07192k5	1992/03/24 11:50	8.067 h	20.39 h	30 m^3	Imatra (coal)
08192n3	1992/03/24 12:00	2.149 d	18.91 h	13 836 m^3	Loviisa
07892n1	1992/03/24 13:10	3.633 h	1.85 d	120 m^3	Loviisa (coal)
06892n4	1992/03/24 13:10	3.617 h	4.50 h	120 m^3	Loviisa
07092n4	1992/03/24 13:10	9.417 h	12.49 h	120 m^3	Loviisa
13692f7	1992/03/24 13:10	1.238 d	17.26 h	120 m^3	Loviisa
04992f8	1992/03/24 14:02	3.233 h	5.45 h	4 137 m^3	Helsinki
20192k1	1992/03/24 16:15	8.217 h	15.80 h	33 m^3	Imatra (coal)
05092f8	1992/03/24 16:20	7.850 h	10.45 h	761 m^3	Imatra
04092f1	1992/03/24 17:26	1.717 h	1.89 h	1 945 m^3	Helsinki
08292n3	1992/03/24 18:13	2.876 d	2.67 d	11 715 m^3	Viitasaari
04192f1	1992/03/24 19:20	5.083 h	1.48 d	70 m^3	Kotka (coal)
13292f5	1992/03/24 19:20	4.817 h	1.49 d	1 020 m^3	Kotka
07492n4	1992/03/24 19:20	2.831 d	2.67 d	1 020 m^3	Kotka
13192f5	1992/03/24 19:32	1.467 h	0.65 h	1 760 m^3	Helsinki
16292f6	1992/03/24 20:50	9.758 d	2.75 d	10 969 m^3	Helsinki
16692n3	1992/03/24 20:50	93.740 d	2.76 d	63 917 m^3	Helsinki
08192n1	1992/03/24 21:03	2.758 d	2.67 d	53 320 m^3	Nurmijärvi
07792n3	1992/03/24 21:09	3.350 h	6.30 h	1 112 m^3	Helsinki
20292k1	1992/03/25 00:05	16.920 h	17.86 h	81 m^3	Imatra (coal)
04292f1	1992/03/25 00:15	1.644 d	23.25 h	895 m^3	Loviisa (coal)
07192n4	1992/03/25 00:15	10.880 h	1.09 d	940 m^3	Loviisa
13392f5	1992/03/25 00:15	1.642 d	19.92 h	1 m^3	Loviisa
07592n4	1992/03/25 00:30	5.362 d	1.94 d	1 849 m^3	Imatra
07992n3	1992/03/25 01:56	9.350 h	7.72 h	7 441 m^3	Helsinki
08092n3	1992/03/25 01:56	17.100 h	18.12 h	7 441 m^3	Helsinki
13692f5	1992/03/25 01:56	5.342 d	1.93 d	7 441 m^3	Helsinki
13992f2	1992/03/25 03:01	1.527 d	23.15 h	150 m^3	Kotka (coal)
15192f6	1992/03/25 03:01	7.867 h	1.04 d	1 737 m^3	Kotka
13592f5	1992/03/25 03:01	2.514 d	2.68 d	1 737 m^3	Kotka
15392f6	1992/03/25 07:57	1.320 d	23.18 h	3 340 m^3	Helsinki
14092f7	1992/03/25 07:57	5.135 d	1.89 d	3 340 m^3	Helsinki
07292n4	1992/03/25 09:14	1.264 d	18.88 h	618 m^3	Helsinki (coal)
07992n1	1992/03/25 11:39	1.163 d	18.93 h	115 m^3	Kotka (coal)
13892f2	1992/03/25 11:39	6.267 h	18.48 h	1 337 m^3	Kotka
15492f6	1992/03/25 11:39	2.154 d	2.67 d	1 337 m^3	Kotka
05292f8	1992/03/25 12:46	5.017 h	18.19 h	357 m^3	Loviisa
05392f8	1992/03/25 12:46	1.123 d	18.98 h	360 m^3	Loviisa (coal)
05592f8	1992/03/25 19:14	1.842 d	2.69 d	4 533 m^3	Imatra
13892f7	1992/03/25 21:19	1.554 d	4.19 h	20 515 m^3	Helsinki
05992f8	1992/03/28 06:12	3.303 d	18.27 h	15 482 m^3	Viitasaari
08892n1	1992/03/28 08:30	4.185 d	20.59 h	15 560 m^3	Imatra
08592n4	1992/03/28 08:30	13.240 d	2.78 d	15 560 m^3	Imatra
07792k5	1992/03/28 08:50	3.200 d	18.90 h	70 300 m^3	Nurmijärvi
14392f2	1992/03/28 09:19	3.217 d	17.86 h	85 023 m^3	Helsinki
17092f6	1992/03/28 09:19	13.240 d	2.74 d	85 023 m^3	Helsinki
16592f6	1992/03/29 01:33	8.613 d	18.05 h	1 628 m^3	Imatra (coal)
15892f2	1992/03/31 08:15	10.280 d	2.74 d	11 643 m^3	Viitasaari
05292f1	1992/03/31 20:15	9.744 d	2.87 d	53 033 m^3	Nurmijärvi
15192f5	1992/03/31 20:30	9.771 d	2.74 d	12 082 m^3	Imatra
15592f7	1992/03/31 21:51	9.717 d	2.74 d	64 289 m^3	Helsinki

Table B.3: Basic information on the test set of 65 Sosnovyy Bor spectra.

Table B.4: Anthropogenic nuclides present in 65 Sosnovyy Bor spectra. The nuclides flagged with y were identified by SHAMAN, those flagged with n were not. *(Table is in 4 parts.)*

Spectrum identifier	^{91}Sr	^{91m}Y	^{95}Zr	^{97}Zr	^{95}Nb	^{97}Nb	^{99}Mo	^{99m}Tc	^{103}Ru	^{105}Rh	^{106}Rh	^{122}Sb	^{125}Sb	^{129m}Te	^{132}Te	^{131}I	^{132}I	^{133}I
13692f2																		
11592f7																		
07292k5																		
15092f7																		
05392f1																		
09392f8																		
04892f1																y		
07392f1																		
06592f8																		
15092f6																y		
13592f7																		
13092f5																		
07792n4																		
13492f7																		
07192k5																		
08192n3			y		y			y	y						y	y	y	y
07892n1																y		
06892n4		y	y	y	y	n	n	y	y	y					n	y	y	y
07092n4	y		y	y	y	n	n	y	y	y					y	y	y	y
13692f7			y	y	y	n	y	y	y	n					y	y	y	y
04992f8																		
20192k1																		
05092f8																		
04092f1																		
08292n3																		
04192f1																y		y
13292f5	y	y	y	y	y	n	y	y	y	y	y		n	y	y	y	y	y
07492n4			y	y	y		y	y	y	y	y				y	y	y	y
13192f5																		
16292f6					y				y							y		
16692n3					y													
08192n1					y			y	y						y	y	n	
07792n3																		
20292k1																		
04292f1						n										y		

Spectrum identifier	^{133}Xe	^{135}Xe	^{134}Cs	^{136}Cs	^{137}Cs	^{140}Ba	^{140}La	^{141}Ce	^{143}Ce	^{144}Ce	^{144}Pr	^{147}Nd	^{224}Ra	^{237}U	^{238}Np	^{239}Np		Total number
13692f2					n													1
11592f7			y		y								y					3
07292k5					y													1
15092f7					y													1
05392f1					y													1
09392f8					y													1
04892f1					y													2
07392f1			y		y													2
06592f8			y		y													2
15092f6					y													2
13592f7																		0
13092f5																		0
07792n4																		0
13492f7																		0
07192k5					y													1
08192n3	y		y		y	n	y	y	y	y		y		y		y		19
07892n1																		1
06892n4		y	y		y	n	y	y	y	y				y		y		23
07092n4	y	n	y		y	y	y	y	y	y		n		y		y		25
13692f7	y	y	y		y	y	y	y	y	y		y		y		y		24
04992f8																		0
20192k1					y													1
05092f8																		0
04092f1			y															1
08292n3					y													1
04192f1					y													3
13292f5	y	y	y	y	y	y	y	y	y	y	y	y		y	y	y		32
07492n4	y		y	y	y	y	y	y	y	y		y		y	y	y		25
13192f5																		0
16292f6			y		y													5
16692n3			y		y							y						4
08192n1			y		y		y	y								y		11
07792n3																		0
20292k1					y													1
04292f1					y													3

Spectrum identifier	^{91}Sr	^{91m}Y	^{95}Zr	^{97}Zr	^{95}Nb	^{97}Nb	^{99}Mo	^{99m}Tc	^{103}Ru	^{105}Rh	^{106}Rh	^{122}Sb	^{125}Sb	^{129m}Te	^{132}Te	^{131}I	^{132}I	^{133}I
07192n4			y	n	y	n	n	y	y	y	y				y	y	y	y
13392f5			y	y	y	n	y	y	y	y	y		y		y	y	y	y
07592n4																		
07992n3			y		y			y	y						y	y	y	
08092n3			y		y	n		y	y						y	y	y	
13692f5			y		y			y	y						y	y	y	
13992f2																		
15192f6					y			y	y						y	y	y	y
13592f5					y		n	y	y			y	n		y	y	y	
15392f6								y										
14092f7								y	y							y		
07292n4																		
07992n1																		
13892f2					y			y	y						n		n	
15492f6			y		y			y	y						y	y	y	
05292f8																		
05392f8																		
05592f8					y			y	y							y		
13892f7																		
05992f8																		
08892n1																		
08592n4					y				y									
07792k5					y			y	y						y	y	n	
14392f2			y		y			y	y						y	y	y	
17092f6			y		y				y							y		
16592f6																		
15892f2																		
05292f1																		
15192f5																		
15592f7					y													
Total	2	2	14	7	24	8	8	21	23	7	4	1	3	1	18	26	18	10

Spectrum identifier	^{133}Xe	^{135}Xe	^{134}Cs	^{136}Cs	^{137}Cs	^{140}Ba	^{140}La	^{141}Ce	^{143}Ce	^{144}Ce	^{144}Pr	^{147}Nd	^{224}Ra	^{237}U	^{238}Np	^{239}Np		Total number
07192n4	y	n	y	n	y	y	y	y	y	y		y		y	y	y		27
13392f5	y		y		y	y	y	y	y	y		y		y	y	y		26
07592n4													y					1
07992n3							y									y		9
08092n3			y		y		y	y						n		y		14
13692f5	y		y		y	y	y	y		y				y		y		16
13992f2					y													1
15192f6			y		y											y		10
13592f5	y		y		y											y		13
15392f6					y													2
14092f7					y													4
07292n4					y								y					2
07992n1					y													1
13892f2								y								y		7
15492f6	y				y		y	y						n		y		13
05292f8																		0
05392f8																		0
05592f8																		4
13892f7																		0
05992f8					y													1
08892n1																		0
08592n4			y		y													4
07792k5			y		y		y	y								n		11
14392f2			y		y		y	y								y		12
17092f6			y		y	n	y	y		y								10
16592f6					y													1
15892f2					y													1
05292f1			y		y								y					3
15192f5					y													1
15592f7			y		y													3
Total	10	5	25	3	47	10	16	16	8	10	1	8	4	11	4	18		393

ISBN 978-951-22-9439-8
ISBN 978-951-22-9440-4 (PDF)
ISSN 1795-2239
ISSN 1795-4584 (PDF)