

MASTER PHYSICS, CONCEPTS AND APPLICATIONS  
École Normale Supérieure de Lyon

Internship 2021–2022  
Antoine Sérandour  
M2 Physics

---

## Data management for radiological characterisation of nuclear waste and materials

---

**Abstract:** Mirion Canberra, a commercial company, provides a range of detectors and softwares for nuclear measurements. The data produced by Genie2k, one of these softwares for gamma spectrometry, are encrypted in data files, which prevents the analysis of interesting features such as trends calculation. The BR3 characterisation group has created an external database to give more flexibility to data management. This study proposes some strategies and their implementations for the improvement of the data management, with adding more data to the process and upgrading the features of the uploading operation. This study calls for further investigations of new data and features to implement, in order to centralise operations, from measurement to regulation report production.

**Key words:** Gamma spectrometry, *in situ* measurement, C# language, SQL, Visual Studio

Internship supervised by:

**BROECKX Wouter**

[wouter.broeckx@sckcen.be](mailto:wouter.broeckx@sckcen.be) / phone (+32) 14 33 21 11

Studiecentrum voor kernenergie ou Centre d'étude de l'énergie nucléaire (SCK CEN)

Boeretang 200

2400 MOL BELGIUM

<https://www.sckcen.be/fr>

sck cen

# Acknowledgement

I would like to thank the whole BR3 team, especially Wouter Broeckx who has always been a good listener. I wish him all the best! I would also like to thank Hasan Moudud for his help and availability! Good luck with your thesis defence! Finally, I would like to thank all those who supported me in my work, especially Arne Van den Brande, Wim De Boeck and Sven Boden. Thank you all and good luck!

# Contents

<b>Introduction</b>	<b>1</b>
<b>1 Characterisation method</b>	<b>2</b>
1.1 Measurement of easy to measure nuclide . . . . .	3
1.2 Calibration and analysis . . . . .	4
1.3 Calculation for difficult to measure nuclides . . . . .	6
<b>2 Current tools used in BR3</b>	<b>6</b>
2.1 Database Organisation . . . . .	7
2.2 Reading and introducing data into the database . . . . .	8
2.2.1 Reading a data file . . . . .	8
2.2.2 Introducing data into the database . . . . .	9
<b>3 Improving the current tools</b>	<b>10</b>
3.1 Inventory of interesting parameters . . . . .	10
3.1.1 List of data parameters . . . . .	10
3.1.2 Focus on efficiency error parameters . . . . .	11
3.1.2.1 Theoretical efficiency calibration . . . . .	12
3.1.2.2 Relevance of efficiency error parameters . . . . .	12
3.2 Database and its implementing tool modification . . . . .	14
3.3 User interface upgrade . . . . .	15
<b>Conclusion</b>	<b>17</b>
<b>A Update procedure</b>	<b>19</b>
<b>B User guide</b>	<b>24</b>

# Acronyms

<b>BR3</b>	Belgian Reactor 3.	<b>LiNDA</b>	Least Common Multiple.
<b>CAM</b>	Configuration Access Method	<b>MSE</b>	Mean Square Error.
<b>CNF</b>	Conjunctive Normal Form.	<b>PWR</b>	Pressurized Water Reactor.
<b>FWHM</b>	Full Width at Half Maximum.	<b>SCK CEN</b>	Studiecentrum voor Kernenergie Centre d'Etude de l'Energie Nucléaire.
<b>HPGe</b>	High Purify Germanium.	<b>SQL</b>	Structured Query Language.
<b>ISOCs</b>	In Situ Object Calibration Software.		

## Introduction

Nuclear power plants have to be dismantled after a certain production time. The main objective of this operation is to remove all nuclear materials from the area. Nuclear power plants produce a high amount of activated and/or contaminated materials, for which management is a key competence in both public health and environment protection.

To guarantee public health, each state created a legal framework to classify waste based on their impact on people and environment, as recommended by the International Atomic Energy Agency [1]. The categories are defined using the activity and the half-life of the atoms in the material. Thresholds were established by each country to separate different types of waste (cf fig 1 from [1]). Thresholds depend on the isotope (called nuclides) that emits the radiation. Therefore, decommissioning a power plant means finding safe, quick and economical ways to sort the waste in categories, which requires determining their composition and concentration (activity) in nuclides. This work is called characterisation.

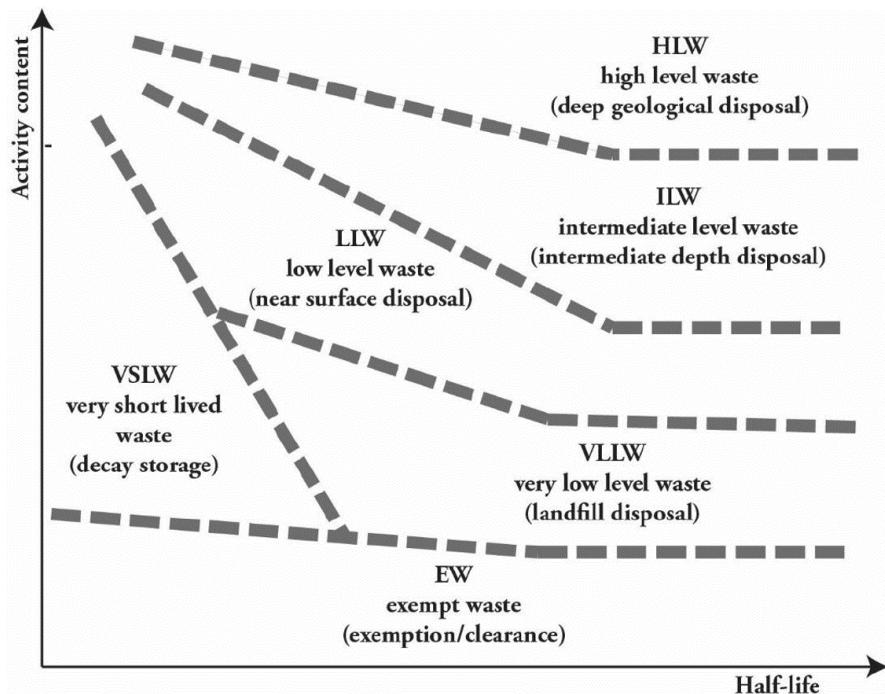


Figure 1: Conceptual illustration of the waste classification scheme from [1].

This internship took place in the BR3 (Belgian reactor 3) in the SCK CEN, the first pressurized water reactor (PWR) in Europe, built in the end of the 1950s. In 1987, it shut down to study the technical and economical feasibility of dismantling. Characterisation activities are a key competence in dismantling, as around 98% of materials from decommissioning are not considered nuclear, because of their low activity ("exempt waste" in the figure 1).

Characterisation of nuclear waste can be performed with the analysis of the radiation emitted by the material, that gives information on the nuclides (this characterisation field is called radiological characterisation). Gamma rays are the easiest radiations to measure and thus are commonly used for characterisation in decommissioning power plant, hence also in the BR3 dismantling. Gamma measurements are also done on site to avoid large volumes of waste analysis in laboratories.

Characterisation methods used in decommissioning reactors produce a lot of data, such as measurement signal, logistic data and analysis results. The analysis of these data is performed by a private company software, that produces encrypted data files. The decommissioning team, operating on the Belgium Reactor 3 (BR3) in the SCK CEN, would like more flexibility in the data management. A database has already been created to contain all these data before this internship. However, the procedure to add data from gamma spectrometry data files to the database is not perfect: more data can

be extracted and the user interface can be improved to enhance the efficiency of the workflow. This work requires the modification of the current database structure and the way to implement data in it. The aim of this internship is thus to upgrade the data management with adding more data to the process and creating more feature to the management.

This report will first describe the characterisation methods to measure the activity and identify nuclides of a nuclear material. Afterwards, the report will address in section 2 the current software tools used in BR3 to perform and process the characterisation measurements. The report will then deal with the improvements implemented to these current tools with this internship.

## 1 Characterisation method

For over 30 years, the BR3 (Belgium Reactor 3) is in the process of being dismantled. At first, the highly contaminated and irradiated materials were removed to reduce radiation levels and avoid contamination effects on less activated material. Today, the remaining materials have low or very low activity, and are mainly composed of infrastructure equipment (wall, floor, etc.). These materials represent a high amount of volume and they are difficult to extract. This is why they are characterised on-site, so these can remain in place if they comply with the activity limits for clearance. Different techniques can be used, the report focused on in situ gamma spectrometry.

Gammaspectrometry measures gamma rays from radioactivity. Radioactivity designates all the transformations of an unstable atomic nucleus into a more stable one. This nucleus will lose energy by emitting a radiation, that can be alpha decay (emission of a helium nucleus), beta decay (emission of an electron or positron), or gamma decay (emission of a gamma ray photon). In a nuclear material, some atoms will emit those radiations. The number of emissions per second is called the activity of the material and is expressed in becquerel. Gamma rays are the easiest rays to detect, because they are not stopped by metal layer (which composes drums for nuclear storage), paint (that can shield contaminated concrete) or centimeters of air (for long distance measurement).

This method is particularly interesting for in situ measurement that consists in placing the gamma ray detector several meter away from the sample (typically a drum) and measure the gamma rays received. Difficult to measure nuclides (alpha or beta emitters) do not reach the detector with their rays, while easy to measure nuclides do and their rays are measured. Afterwards, using reference nuclides and the measurement set up, the spectrum is calibrated and analysed to obtain activities or detection limits of all easy to measure nuclides. Finally, a calculation method called "scaling factor" gives activities for difficult to measure nuclides. The figure 2 summarises the whole process. This procedure is detailed in this section.

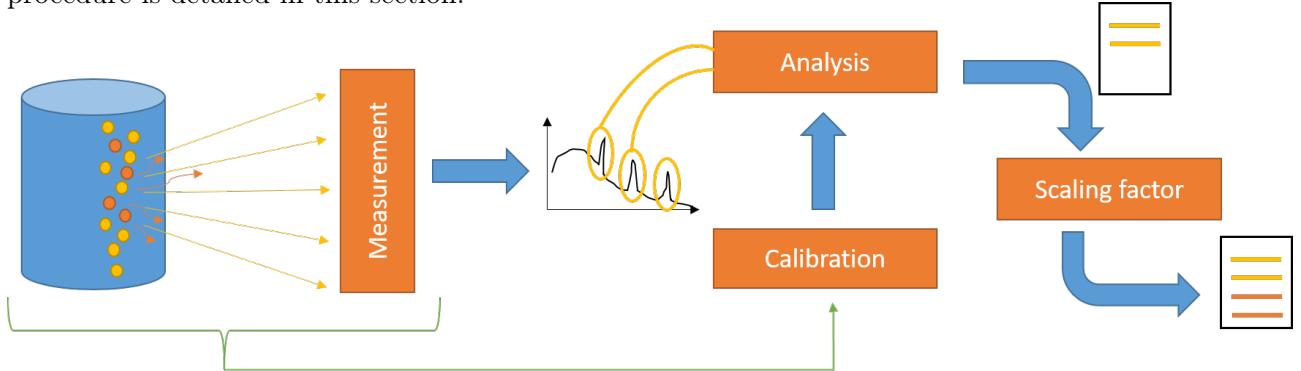


Figure 2: Scheme of the characterisation process. Yellow color designates easy to measure (gamma emitters) nuclides in the waste, and orange color designates difficult to measure nuclides. In the end, the calculation gives the activity and composition of both easy and difficult to measure nuclides.

Gamma measurements are explained in subsection 1, followed by subsection 2 that gives an overview of the analysis and calibration techniques to obtain activity values. This section ends dealing with the scaling factor method to obtain activities of nuclides that do not produce gamma rays.

## 1.1 Measurement of easy to measure nuclide

These measurements of gamma rays are performed by high purity germanium spectrometer (HPGe spectrometer), a photography is shown in figure 3.



Figure 3: Picture of the detector setup. The beige and red tank contains liquid nitrogen, in front of which is the germanium piece inside the lead collimators, grey-blue on the picture. The sample measured is inside the bottle in front of the collimators.

The detector operation is based on excitation of matter with the gamma photons emitted by nuclides. There are three types of interactions between photons and matter: photoelectric effect, Compton scattering and pair production. The photoelectric effect designates the emission of an electron by an atom that absorbs a photon, the photon gives all its energy to the electron (ionisation and kinetic energy). The Compton scattering corresponds to the emission of an electron while a photon scatters through the electronic cloud, a part of the energy of the photon is transferred to the electron. The pair creation refers to the disintegration of a photon into an electron-positron pair, the photon gives all its energy to these particles (mass and kinetic energy). The proportion of these phenomena is related to the energy of the photon. Typical gamma rays emitted by nuclides have an energy in [20 keV, 2 MeV], so the preponderant interaction is the Compton scattering.

The detector used in BR3 is equipped with a germanium crystal that is cooled with liquid nitrogen to increase its semiconductor properties. When a gamma ray gets into the germanium material, the photon excites one electron by Compton effect. This electron receives a certain amount of energy and reaches the conduction band of the germanium. It also creates a hole, that can be really deep in the valence band as the gamma ray was highly energetic. The hole and the electron redistribute then the energy they received to create other electron-hole pairs. Therefore, all the resulting electron/hole pair have an energy slightly above/below the band gap. The number of pairs is thus proportional to the energy of the incident photon. An electric field is applied, so that the electron/holes are dragged toward the cathode/anode. The signal is multiplied and the total charge gives the energy of the incident photon. This operation is described in [2].

The HPGe spectrometer counts the radiations it receives and measures the voltage each radiation creates. To return a spectrum, the voltage scale is divided into channels. The raw spectrum thus gives the count rate of received radiations with respect to the voltage. Calibration techniques, that are discussed in the next section, will provide the decay rate with respect to the energy. A lead collimator is used in front of the detector to focus on radiation coming from the sample and avoid background

noise.

Figure 4 shows an example of a spectrum obtain using such a setup. The typical energy scale of the gamma rays is from 20 keV to 2 MeV. The spectrum is composed of several peaks at discrete energies, and a background signal.

The peak location is the signature of the nuclides: each nuclide, when decaying via gamma emission, loses a fixed amount of energy. The energy of the measured ray correspond to this loss. As one nuclide can decay with different radiations (alpha, beta, gamma, at different energy), several peaks can be observed for one nuclide. This property recalls a similar phenomenon of the UV visible emission spectrum: each atoms has energy levels for its orbiting electrons, a transition for an electron from a level to another leads to the emission or absorption of a photon with a given energy. The peak energy of the emission spectrum corresponds to the atom signature. In the same way, gamma spectrometry studies emission spectra to analyse peaks and find the related nuclides. For gamma spectrometry, the relative intensity of these peak is related to the branching ratio: the probability for an excited nuclide to decay in a particular way.

However, in a UV-visible emission spectrum, there is no background signal. The gamma background is due to the Compton scattering in the germanium: before being absorbed, gamma rays can scatter and their energies are spread, which creates a background signal. In UV-visible emission, photons interact mainly with photoelectric effect, where the photon energy is fully absorbed and thus no background is generated.

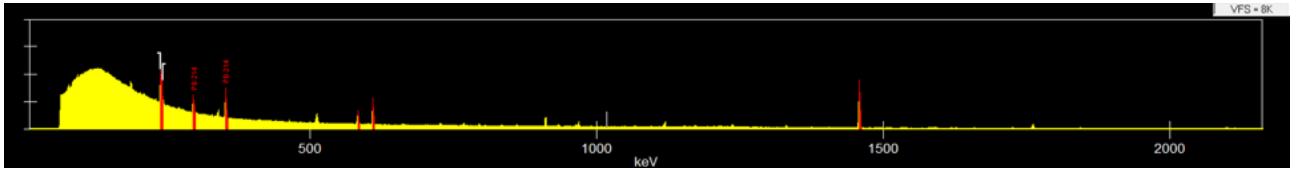


Figure 4: Typical gamma spectrum of a wall waste sample. Peaks appear above the background and give the signature of the waste nuclide composition.

## 1.2 Calibration and analysis

The analysis of the spectrum depends on the objective of the measurement. The main aim is often to determine the radioactivity of important gamma emitters in the material. The analysis is composed of several steps that are performed numerically, an overview is given in figure 5. All the following details are given in [2].

The first step to analyse a spectrum is the localisation of the peaks. Several numerical methods can be used. The software Genie2k from the private company Canberra, that is the one used in the BR3 team to acquire and analyse the spectrum, is based upon the Mariscotti algorithms using the second difference method. This procedure calculates the second derivative of the spectrum using second difference, and when the algorithm meets a peak, the second derivative achieves its minimum at the centroid of the peak. Whichever method is used, the algorithm needs a sensitivity parameter to distinguish between peaks and noise. This criteria should be set up properly to avoid too many reported peaks with some spurious ones as well as too few peaks with some real peaks ignored. Once the peak is localised, some tests on the shape of the peak can be performed.

The next step is the calibration of the raw spectrum: the spectrometer does not provide directly the activity with the energy, but counts the radiations that creates a voltage. Calibration is performed in three parts:

- Energy calibration - the relationship between channels and energy ;
- Peak width calibration - the variation of peak width with energy ;
- Efficiency calibration - the relationship between number of counts and disintegration rate (all emitted radiations do not reach the detector crystal)

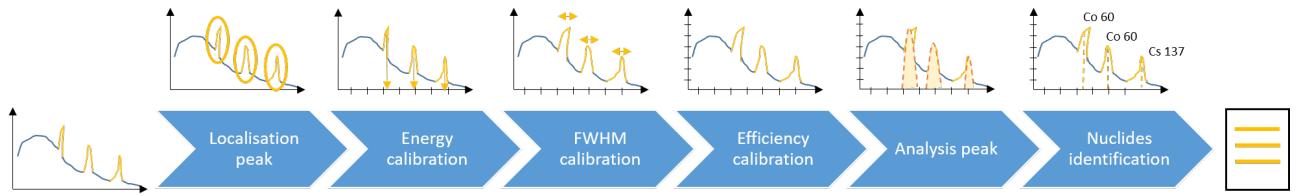


Figure 5: Analysis sequence of a spectrum: from a raw spectrum, the analysis is able to give the composition and the concentration in nuclides in the sample.

Calibration is based on the initial measurement of a designed sample with known nuclides. These nuclides are often fission products that have a particular signature. The nuclides used in BR3 are Am 241 and Eu 152.

During energy calibration, a range of peaks are used as reference using the library, and the measured centroids of these peaks are compared with the expected centroids energy of these peaks. A linear or quadratic fit is performed giving the relation between the channel number and the energy.

The peak width calibration has the same procedure than the energy calibration. The full width at half maximum (FWHM) measured from the reference source is compared to the known FWHM. This calibration seems to be optional, but it increases a lot the accuracy on the shape of the peak, that is crucial in determining activities.

Finally, the efficiency calibration is performed. In gamma spectrometry, the efficiency is defined as the relation between a peak area and the amount of radioactivity the nuclide characterized by this peak represents. This efficiency can thus be written as:

$$\epsilon = \frac{R}{SP}, \quad (1)$$

with  $\epsilon$  the efficiency,  $S$  the actual activity of the sample,  $R$  the count measured and  $P$  the branching ratio, ie the probability for a nuclide to emit a gamma ray. As a reminder, all the parameters in the previous equation are energy dependant, this is why  $\epsilon$  is called an efficiency curve.

The parameters that can affect the efficiency are numerous. For example, the source is not a point source, and the detector is placed at a distance of the source, so that all gamma rays coming from the source are not caught by the detector. Either they are missing the detector (due to the distance to the source or the height of the source), or they are absorbed by the sample itself (due to the density of the source).

Therefore, the efficiency calibration cannot be based on the measurement of the known sample, as too parameters are involved. The calculation of the efficiency curve is mainly done using modelling, the geometry of the experiment and some properties of the sample such as density. The model indicates the efficiency for some reference energies and a polynomial fit gives the efficiency curve. The reference energy and the order of the polynomial fit depend on the measurement setup.

After calibration, the analysis performs the characterisation of the different peaks in the spectrum. For each significant peak the energy, the height, the full width at half maximum FWHM and the area are calculated, with an uncertainty on every of these parameters. This analysis requires to have a good resolution for the peaks, to deconvolute multiplets and to make corrections.

The final step of the analysis is the nuclide identification. The peak parameters (shape, centroid energy, ratio with other peaks) give information on the initial nuclide. A nuclide library contains many theoretical peak parameters and the software identifies each measured peak to a theoretical peak and thus a nuclide. The used library has a huge influence on the resulting identification: a too short list will fail linking the observed peak to the nuclides, whereas a too long list will lead to mismatch in the linking (many possible nuclides for one peak energy).

At this point, all the peaks have been analysed and are related to a nuclide. The area of the peaks give the counts of the gamma rays emitted by one or more nuclides at that peak energy. Using the time of measurement, the software gets the count rate for all the peaks, and then attributes this activity to a nuclide. As one nuclide can lead to several peaks, the software calculates several activities for one nuclide, and calculates a mean weighted with the uncertainties on the activities.

### 1.3 Calculation for difficult to measure nuclides

This analysis described so far gives the activities of the nuclides measured with gamma spectrometry, but other nuclides are not detectable with this method: some nuclides decay with only alpha or beta rays or have a tiny branching ratio, some nuclides have a very light activity and the peaks are hidden in the background, etc. To address this problem, the scaling factor method is often used.

The scaling factor method is a technique used to determine the activities of difficult to measure nuclides with some key nuclides, that are easy to measure [3]. At first, a preliminary study analyses the nuclides (easy and difficult to measure) in a typical sample at a reference time, using laboratory measurements. Some nuclides are chosen to be key nuclides, and for each nuclide detected the ratio of its activity over the one of the reference nuclide is calculated. These ratios will evolve with time, as the half lifes of nuclides are not equal. Using nuclear exponential decay and the branching ratio, the evolution of the ratios are determined. The scaling factor method consist in calculating the activities of difficult to measure nuclides by multiplying the related ratio with the activity of the key nuclide.

This method minimizes the use of destructive assays in laboratory, a time consuming and expensive process that exposes the operational personnel to a not necessary dose. However, the choice of the key nuclide and the typical sample is decisive. Only a limited number of nuclides can be used as key nuclides: it should be an easy to measure, with a great concentration and with a significantly long enough half-live. The typical key nuclides used are Co-60, Cs-137 and Sb-125. The typical sample used should correspond to the material that contains the higher concentration in radionuclides. In BR3, these samples were taken from the 'crud', that is the sediment in the primary water of the reactor. This sediment is highly contaminated, and particularly interesting since it gives information on the nuclides that have a good solubility and mobility. These nuclides have spread during the operation of the power plant, contaminating the whole primary circuit, and thus the scaling factors refers to this contamination.

The gammaspectrometry and the scaling factor method provide a way to determine the nuclide composition and activity of a material. This methodology requires softwares to perform the calculation and numerical tools to manage the data. This internship was exploring this data management. The current data management status is described in the following section.

## 2 Current tools used in BR3

The measurement and analysis process uses several software tools. In the nuclear industry, the private company Canberra is one of the main suppliers of detectors and softwares. The BR3 team uses their HGe detector with the Genie2k software to perform the measurement and the analysis process. This software puts every of its resulting data in one file with the extension CNF (Conjunctive Normal Form), that contains around 3700 parameters.

The measurement data and the analysis results are thus stored in this encrypted CNF file, that can only be read by a Canberra software such as Genie2k. These data can be displayed in the Genie2k user interface which is not a flexible tool to access, process and analyse these data. This is why the team created an SQL database in which results of a measurement and analysis are stored. This SQL database is manipulated using SQL Server Management from Microsoft. Some of the data (around 90 parameters) are extracted from the CNF files using an in-house developed C# application called CaTo1800. This application was part of a project (LiNDA) created by a laboratory team, to centralise all the characterisation operations (measurement, analysis and data management) in one big application. CaTo1800 is thus a sub application of this project, that transfers data from CNF files into the database. The following section explains the current data management of the characterisation team, that is illustrated in the figure 6.

This section begins with an overview of the database structure detailing the parameters types, which is followed by a description of the CaTo1800 script in C# regarding the extraction and uploading of data from CNF to the database.

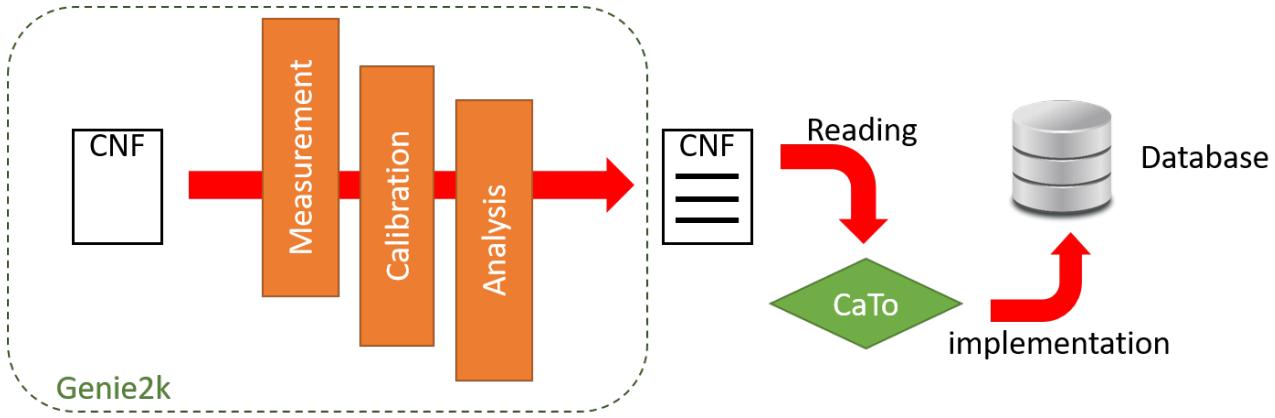


Figure 6: Scheme of the data management process. Genie2k produces data encrypted files called CNF, and CaTo1800 reads and implements the data in the database.

## 2.1 Database Organisation

The current SQL database used to store the CNF data files is called ISOCSBR3 (In Situ Object Calibration Software Belgium Reactor 3). It is composed of 5 tables (Spectrum, ResultsLines, ResultsNDA, Peak and ResultsNuclides), organised as shown in figure 7. 90 parameters extracted from the CNF files are distributed across these 5 tables.

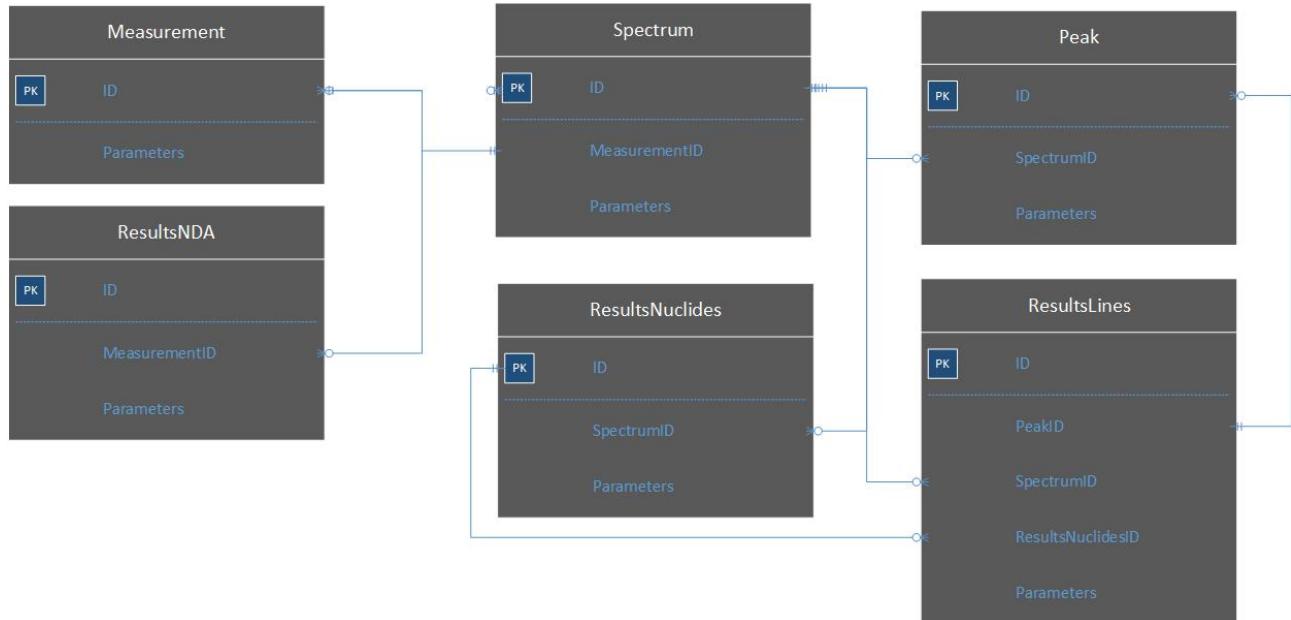


Figure 7: ISOCSBR3 initial database structure with six tables containing 90 parameters.

Each table contains data of the analysis of the spectrum:

- Table Measurement: each line refers to a sample with its typical data (weight, volume, ID number etc.)
- Table Spectrum: each line refers to a measurement of a spectrum with data concerning the setup of the experiment (measurement date, the detector name, collimator used, etc.) and is related to one sample with the MeasurementID.
- Table Peak: contains all the data about the peaks located in one spectrum (energy, area, full width at half maximum, etc.). The lines are related to the spectrum table with the SpectrumID.

- Table ResultsLines: for each nuclide listed in the library of Genie2k, informations about all the expected peaks are calculated (energy, measured activities, minimum detectable activities, uncertainties, etc.) and Genie2k makes links between expected peaks and measured peaks.
- Table REsultsNuclides: each line represents the total activity of a nuclide listed in the library (activities, minimal detectable activities, uncertainties, etc.)
- Table ResultsNDA: summary table of nuclides and their activities in every spectrum.

This SQL database allows more flexibility in data management and analysis than a user interface. Moreover, Canberra gives some tools to access these data: a program executable via command prompt, that can extract a parameter from one CNF file, but this method is time consuming for extracting a lot of data. Canberra also proposes a C# library able to extract parameters from a CNF file. The CaTo1800 software was created at SCK by Wim De Boeck to read, extract and import data from CNF files to a database, using this library.

## 2.2 Reading and introducing data into the database

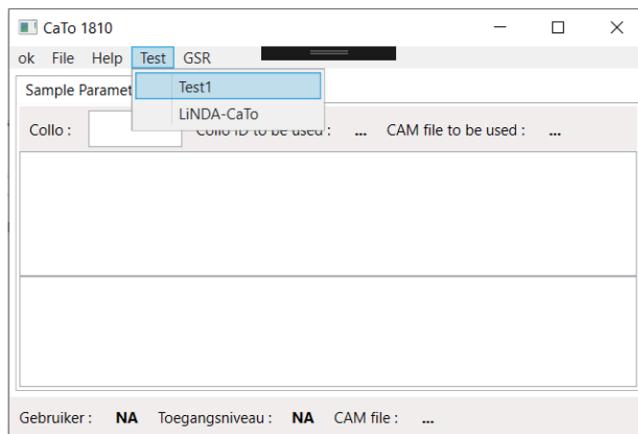


Figure 8: Initial CaTo1800 interface. "File" and "Test" menu items are the only relevant tools for reading and uploading data to database.

CaTo1800 is an application (cf fig 8) composed of several scripts and an user interface tool. It basically realises two operations: reading CNF files and transfer this data in the database, that are described in the next subsections.

### 2.2.1 Reading a data file

CaTo1800 can open a CNF file, read and store temporarily the data in an object called a class.

To read CNF files, Canberra elaborated a C# library to access to the data in the CNF file. The shape of the function is standardized: CAM\_type\_name . Where "CAM" (Configuration Access Method) is the general name to begin all access functions, "name" is the name of the parameter returned and "type" specifies the returned type of the function. This type is designed by a letter (F: a float, X: a date or time, T: a character string). The part of the code dealing with this operation is shown in figure 9. The code defines a public method that will first opens the CNF file and then gets the value of each of the previously defined parameters. The keyword try and catch refers to the case of an error while reading the file.

After reading the data, they are stored in the class. The properties of the class that contains these data are of two types: private properties (the data) and public properties (the functions to add or rename a parameter).

```

1562     public Boolean ReadAcquisitionParamters()
1563     {
1564         Boolean IsReadSampleParametersSucces;
1565         IsReadSampleParametersSucces = true;
1566         try
1567         {
1568             m_CAM.Open(m_FileName, CanberraDataAccessLib.OpenMode.dReadWrite, 4096);
1569
1570             try
1571             {
1572                 m_CAM_X_ASTIME =
1573                     System.Convert.ToDateTime((m_CAM.get_Param(CanberraDataAccessLib.ParamCodes.CAM_X_ASTIME, 1, 1)));
1574             }
1575             catch (Exception)
1576             {
1577                 //m_CAM_X_ASTIME = DateTime.
1578             }
1579         }

```

Figure 9: Part of the CaTo1800 script on reading parameters.

The private properties code is shown in the figure 10. The parameters are set as private properties of the class and the names used are following the terminology of the constructor, with m\_ added as a prefix. This code is iterated to describe all the parameters of interest in a CNF file.

```

93     private string m_CAM_T_CPNNAME; // Configuration physical name (32 characters) (p.487)
94     private string m_CAM_T_ACQMOD; // Acquisition mode (PHA+, PHA-, MCS+)
95     private float m_CAM_X_PREAL; // preset real time (p. 488)

```

Figure 10: Part of the CaTo1800 script on private properties.

The public definition is shown in figure 11. The code defines as a public property (callable from elsewhere) the function giving access to the data stored in the class: the method gets the data value stored, or sets it to a new value.

```

144     public string CAM_T_CPNNAME
145     {
146         get { return m_CAM_T_CPNNAME; }
147         set { m_CAM_T_CPNNAME = value; }
148     }
149
150     public float CAM_X_PREAL
151     {
152         get { return m_CAM_X_PREAL; }
153         set { m_CAM_X_PREAL = value; }
154     }

```

Figure 11: Part of the CaTo1800 script on public properties.

### 2.2.2 Introducing data into the database

The second operation of CaTo1800 is the introduction of the data into the database. The program creates a new operator (named ta) which is able to modify a table of the SQL database. A stored procedure, that is an SQL request applied to the database and callable by a script, is used to add a line to a particular table with some values. The script is shown in figure 12.

CaTo1800 has a lot of benefits: firstly, it is easy to handle for operators and it extracts in a short time a lot of data. Secondly, it extracts separately the parameters, and thus can adapt the implementation in the database with its structure. Thirdly, while it imports data in the ResultsLines, CaTo1800 assigns the peaks to every characterized nuclides. This relation between nuclides and peaks is not directly implemented in the CNF files, as data are separately callable with CAM request.

```

96     void WriteToDB(clsCAM objC)
97     {
98         ISOCsBR3database.ISOCsBR3dbTableAdapters.QueriesTableAdapter ta;
99         ta = new ISOCsBR3database.ISOCsBR3dbTableAdapters.QueriesTableAdapter();
100
101        //NDAdatabase.NDAdbTableAdapters.QueriesTableAdapter ta = new NDAdatabase.NDAdbTableAdapters.QueriesTableAdapter();
102        int? Id = new int();
103
104        for (int i = 0; i < objC.CAM_CLS_PEAK; i++)
105        {
106            Id = new int();
107
108            ta.uspPeakInsert(m_SpectrumID, objC.CAM_CLS_PEAK, objC.CAM_F_PSCENTRD[i], objC.CAM_F_PSDCENTRD[i],
109                            objC.CAM_F_PSENERGY[i], objC.CAM_F_PSDENERGY[i], objC.CAM_L_PSLEFT[i], objC.CAM_L_PSWIDTH[i],
110                            objC.CAM_F_PSDBACKGND[i], objC.CAM_F_PSDBACK[i], objC.CAM_F_PSFWHM[i], objC.CAM_F_PSAREA[i],
111                            objC.CAM_F_PSDAREA[i], objC.CAM_F_PSORIGAREA[i], objC.CAM_F_PSORIGERR[i], objC.CAM_F_PSLCCENT[i],
112                            objC.CAM_F_PSDLOCCENT[i], objC.CAM_F_PSSIGNIF[i], objC.CAM_F_PSGROSS[i], objC.CAM_F_PSAMBBACK[i],
113                            objC.CAM_F_PSDAMMBACK[i], objC.CAM_L_PSPKNOWN[i], objC.CAM_L_PSPPFIT[i], objC.CAM_L_PSPMULT[i],
114                            ref Id);
115            objC.PeakID.Add(Id.Value);
116        }
117    }
118 }

```

Figure 12: Part of the CaTo1800 script on writing to database

### 3 Improving the current tools

CaTo1800 is a highly interesting tool, especially because of the opportunities it opens. Created by a laboratory team, this application is the link between the data files and the database. This tool could allow the team to become more independent from the Canberra analysis, focusing more on data collection and data analysis, which is also the aim of the LiNDA project that created CaTo1800. To achieve such objectives, two main strategies are followed: implementing more data to the process by adding parameters to the database, and enhancing the features of CaTo1800 to increase the flexibility of the data by reading the database.

The improvements of the data management are addressed in the following order: first, by creating an inventory of additional parameters of interest in a CNF file. Second, by extracting and implement these parameters of interest from the CNF file into the database. Third, by improving the user interface to have an easier access to the database from CaTo1800, that cannot read the database for now.

#### 3.1 Inventory of interesting parameters

The data management needed more flexibility to store data from laboratory and in situ. The first objective of the internship was thus to elaborate the list of parameters that are also relevant for in situ gamma spectrometry measurements, as CaTo1800 was first designed by a laboratory team. The CNF files contains around 3700 parameters that are not all of interest. To build this inventory, we first had to detail and clarify all the parameters. Afterwards, some error in efficiency calibration parameters were not clearly understood and we performed an analysis to decide whether the parameters were interesting or not for characterisation. This subsection first deals with providing an outlook of all parameters, and then addresses the meaning of calibration error parameters.

##### 3.1.1 List of data parameters

To elaborate a list of all the parameters contained in a CNF file, we proposed an excel sheet (cf fig 13) that would contain all the information describing each parameter, such as the name of the CAM request (that is the identifier of a parameter), a short description, the type of data (date, float, etc.), and some examples from several CNF files. We also looked for parameters that would provide different results on our examples to highlight parameters that have high variability.

The list of all CAM request with their description were given in a extensive Canberra pdf document and were extracted with a python 3 script. This file provided also a table of content to sort the

parameters. The example values were obtained with another python 3 script using the executable program from Canberra that is able to ask a parameter from a CNF file via the command prompt.

The team selected a range of 80 parameters to add to the 90 initial ones, and we can highlight some particular types of parameters:

- Acquisition and sample parameters: give general information about the measurement (preset time of measurement, name of the detector,...).
- Calibrations parameters: (in energy, FWHM and efficiency) give the values of the coefficients, the error, the order of the fit, etc.
- Flags parameters: bit parameters (1 or 0) that give information whether an action (measurement, analysis, calibration) was performed or not.
- Record parameters: contain several values for one request (the error of FWHM for each peak for instance).

1	Col	subcontent	subsub	CAM request	description	type parameter	value for 114-0	value of 14-0	extra	Is common	Column1	BR3-1	BR3-2	BR3-3	Emme	Filter
24	Acquis Common Parameters	CAM_X_PLIVE		Preset live time	Delta date/time par	10000 secs.	0 secs.	y	FALSE		60	100	10	500	10	
176	Acquis Common Tabular Param	CAM_X_EREAL		Elapsed real time (see note 2 on page 511)	Delta date/time par	3002.36 secs.	60238.9 secy		FALSE		600	1000	100	5002.5	100	
177	Acquis Common Tabular Param	CAM_X_ELIVE		Elapsed live time (see note 2 on page 511)	Delta date/time par	3001.29 secs.	60172.1 secy		FALSE		60	100	10	500	10	
372	Acquis Record Param	Multichan CAM_T_MCAID		MCA serial number (8 characters)	String parameter	13000711	13000711 y		FALSE		1E+07	1E+07	1E+07	1,3E+07	1E+07	
384	Acquis Record Param	Energy Cal CAM_X_ECALTIME		Energy calibration time	Absolute date/time	2021-03-29 13	2021-03-29 13 y		FALSE		2020-06	2021-03	2021-1	2020-06-	2020-0	
385	Acquis Record Param	Energy Cal CAM_I_ECALTERMS		Number of terms in energy calibration polync Long int parameter	Long int parameter	2	2 y		TRUE		2	2	2	2	2	
386	Acquis Record Param	Energy Cal CAM_F_ECOFFSET		Energy calibration offset	Float parameter	-0,055861	-0,055861 y		FALSE		0	-0,056	-0,06	0	0	
387	Acquis Record Param	Energy Cal CAM_F_ESCLOPE		Energy calibration slope	Float parameter	0,250359	0,250359 y		FALSE		0,2502	0,2504	0,25	0,2502	0,2502	
388	Acquis Record Param	Energy Cal CAM_F_EQUAD		Energy calibration quadratic term	Float parameter	0	0 y		TRUE		0	0	0	0	0	
389	Acquis Record Param	Energy Cal CAM_F_ECALFAC1		Energy calibration cubic term	Float parameter	0	0 y		TRUE		0	0	0	0	0	
390	Acquis Record Param	Energy Cal CAM_F_ECALFAC2		Energy calibration 4th order term	Float parameter	0	0 y		TRUE		0	0	0	0	0	
391	Acquis Record Param	Energy Cal CAM_F_ECALFAC3		Energy calibration 5th order term	Float parameter	0	0 y		TRUE		0	0	0	0	0	
393	Acquis Record Param	Energy Cal CAM_F_ECALCHISQ		Energy calibration Chi square	Float parameter	256,095184	256,095184 y		FALSE		1	256,1	256,1	1	1	
401	Acquis Record Param	Peak Shap CAM_T_FWCAALTYPE		FWHM calibration type (8 characters)	String parameter	SQRT	SQRT	y	TRUE		SQRT	SQRT	SORT	SQRT	SQRT	
402	Acquis Record Param	Peak Shap CAM_X_SCALTIME		Shape calibration time	Absolute date/time	2021-03-29 13	2021-03-29 13 y		FALSE		2020-06	2021-03	2021-1	2020-06-	2020-0	
403	Acquis Record Param	Peak Shap CAM_F_FWHMOFF		FWHM offset: square root equation	Float parameter	0,261438	0,261438 y		FALSE		1,0787	0,2614	0,261	1,07871	1,0787	
404	Acquis Record Param	Peak Shap CAM_F_FWHMSLOPE		FWHM slope: square root equation	Float parameter	0,033482	0,033482 y		FALSE		0,0276	0,0335	0,033	0,02763	0,0276	
405	Acquis Record Param	Peak Shap CAM_F_FWHMHISQ		Square-root FWHM calibration Chi square	Float parameter	7,687204	7,687204 y		FALSE		1	7,6872	7,687	1	1	
452	Acquis Record Param	Miscellane CAM_T_DETNAME		Detector name (16 characters)	String parameter	B15061	B15061 y		TRUE		B15061	B15061	B1506	B15061	B15061	
460	Acquis Record Param	Miscellane CAM_T_DETTYPE		Detector type (ie, NaI, HPGe, Alpha, etc.) (8ch)	String parameter	Ge	Ge	y	TRUE		Ge	Ge	Ge	Ge	Ge	
461	Acquis Record Param	Miscellane CAM_T_DETENGTYP		Detector type: Low-energy vs. Normal (8char)	String parameter	Normal	Normal	y	TRUE		Normal	Normal	Norm	Normal	Normal	
469	Sampl Common Parameters	CAM_T_STITLE		Sample title (64 characters)	String parameter	114-00-K1_01	141-04-K2_01 y		FALSE		B3-14-C	B3-21-	B3-2	Emmer_	Filter_E	
470	Sampl Common Parameters	CAM_T_SIDENT		Sample identification (16 characters)	String parameter	114-00-K1-01	141-04-K2-01 y		FALSE		B3-14-C	B3-21-	B3-2	Emme_2	filter_E	
471	Sampl Common Parameters	CAM_T_STYPE		Sample type (16 characters)	String parameter	vloer	muur	y	FALSE		SLAB	1m <sup>3</sup>	1m <sup>3</sup>	betonpui	filter	
472	Sampl Common Parameters	CAM_T_SGEOMTRY		Sample geometry (16 characters)	String parameter	vloer	muur	y	FALSE		SLAB	1m <sup>3</sup>	1m <sup>3</sup>	betonpui	filter	
492	Sampl Common Parameters	CAM_T_SDDESCN		Sample description n (64 characters), where r	String parameter	(B15061,afstand:5(B15061,afstand:y	(B15061,afstand:5(B15061,afstand:y		FALSE		['affsch	['B1506	['B15	['machie	['affsch	
615	Geom Record Parameters	CAM_X_DCALTIME		Detector efficiency calibration time	Absolute date/time	2021-04-02 8	2021-12-07 13 y		FALSE		2020-11	2021-08	2021-1	2020-07-	2020-1	

Figure 13: Excel sheet list parameters, giving for each parameter a content location (3 first columns), its name (4th column), its description (5th column), its datatype (6th column), two example data (7th and 8th columns), the selection of the team ("y" means yes to extraction, 9th column), the "IsCommon" property (are all the example data equal, 10th column), and other example data.

Most of the parameters were clearly understood by the team, but some particular data on error in efficiency calibration were not clear and not fully described in the Canberra information sheet. Before adding it to the data management, an analysis was performed to investigate the relevance of these parameters.

### 3.1.2 Focus on efficiency error parameters

The BR3 decommissioning team is specialized in radiological characterisation using ISOCS measurements, as it is an efficient way to characterize large quantities of activated materials. As explained in the first section, efficiency calibration has a great impact in the determination of the activities, and thus the parameters referring to the efficiency calibration are of high interest. One of the objectives in the future is to investigate relationships between models (shape, density, distance to the source,etc.) and the efficiencies.

Understanding the efficiency parameters and incorporate them in the database is thus of high interest and this is why we investigate efficiency and efficiency error parameters. The first subsection deals with theoretical determination of efficiency coefficients and standard deviation. Afterwards an analysis of the relevance of these parameters is provided.

### 3.1.2.1 Theoretical efficiency calibration

Efficiency calibration consists in having a model that calculates  $k$  efficiencies  $\{\epsilon_j\}_{1 \leq j \leq k}$  at reference energies  $\{E_j\}_{1 \leq j \leq k}$  and then fitting it to obtain the efficiency curve. The type of the calibration can be linear, logarithmic or empirical, depending on the shape of the efficiency curve. Calling  $n$  the order of the polynomial, we have:

$$\log(\epsilon) = \sum_{i=0}^n c_i \left[ \frac{1}{E} \right]^i \quad \text{for linear calibration} \quad (2)$$

$$\ln(\epsilon) = \sum_{i=0}^n c_i [\ln(E)]^i \quad \text{for logarithmic calibration} \quad (3)$$

$$\ln(\epsilon) = \sum_{i=0}^n c_i \left[ \ln\left(\frac{(E_{\max} + E_{\min})/2}{E}\right) \right]^i \quad \text{for empirical calibration} \quad (4)$$

$E_{\max}$  and  $E_{\min}$  largest and smallest calibration energies

For ISOCS measurement at BR3, logarithmic calibration is performed, and thus the analysis focuses on that type of calibration. To calculate the  $c_i$ , Genie2k uses least square method using matrix formalism:

$$Y = X \cdot C + \eta \quad (5)$$

$$\text{with } Y = \begin{bmatrix} \ln(\epsilon_1) \\ \vdots \\ \ln(\epsilon_k) \end{bmatrix}; X = \begin{bmatrix} 1 & \ln(E_1) & \cdots & \ln(E_1)^n \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \ln(E_k) & \cdots & \ln(E_k)^n \end{bmatrix}; C = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} \text{ and } \eta = \begin{bmatrix} \eta_1 \\ \vdots \\ \eta_k \end{bmatrix}$$

The  $C$  matrix is determined so that  $\eta^t \eta$  is minimal, where  $\eta^t$  is the transposed matrix of  $\eta$ . Theoretical work on regression [4] gives the expression of the  $C$ :

$$C = (X^t X)^{-1} X^t Y \quad (6)$$

Considering now  $\eta$  as a vector of random variables, a good estimator of the correlation matrix is given by:

$$\hat{s}^2(\eta) = MSE[I - X(X^t X)^{-1} X^t] \quad (7)$$

with  $MSE = \frac{1}{k-2}(Y - XC)^t(Y - XC)$ . Finally, an estimator of standard deviation of the efficiency fit is given with:

$$\hat{s}^2(Y_E) = MSE[X_E^t (X^t X)^{-1} X_E] \quad (8)$$

with  $Y_E = [E]$  and  $X_E^t = [1 \quad \ln(E) \quad \dots \quad \ln(E)^n]$ .

This matrix formula can be written as  $\hat{s}^2(\epsilon_E) = \sum_{i=0}^{2n} d_i \ln(E)^i$ . Among calibration parameters, Canberra gives the list of the  $c_i$  and the list of the  $d_i$ . Knowing the meaning of these data, their relevance can be addressed.

### 3.1.2.2 Relevance of efficiency error parameters

Using other parameters provided by Genie2k, we can understand what role these parameters play in the analysis.

First the efficiency fit and standard error are plotted and compared with manually calculated fit (using matrix method with numpy library in Python 3). The points used for calibration were added as well. All these data are shown in the figure 14. They are perfectly overlapping for the efficiency fit, but not for the error. This indicates that the assumed error parameters (called Herrmat in the Canberra pdf document) is wrongfully interpreted and do not represent the error as plotted below. However, the values given by manually calculated fit are close to error points. These error points are given with the efficiency point while building the model. We used the manual fit to evaluate the relevance of the parameter, without knowing how it can provide correct values.

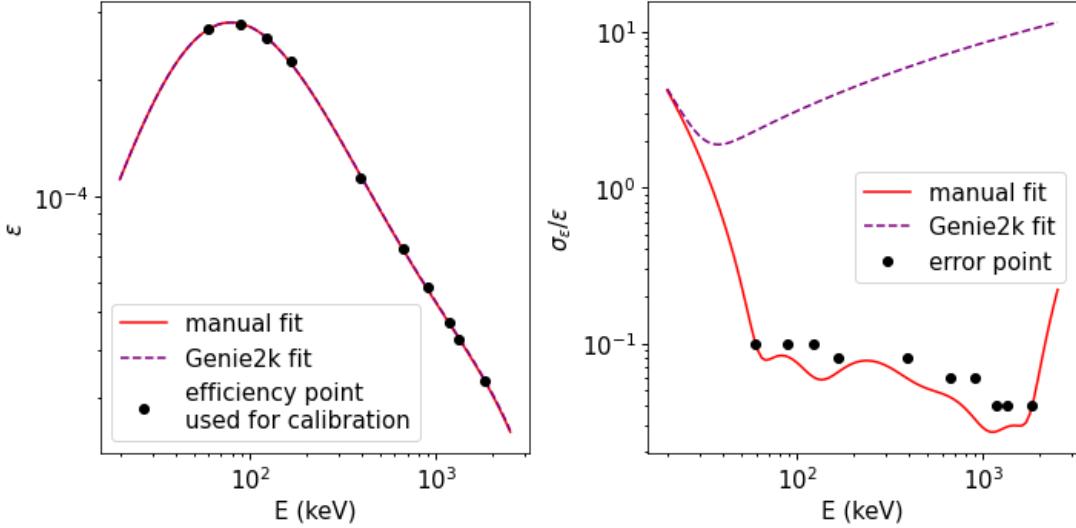


Figure 14: Comparison between Genie2k calculation and manual fit calculation, for efficiency (left) and its standard deviation (right). Genie2k and manual fits overlap for efficiency but not for standard deviation. Given values of Genie2k fit, the parameter has not been fully understood.

To evaluate the relevance of these parameters, we focused on their use in the characterisation method. As explained in subsection 1.2, Genie2k expects some nuclides in the material, that emit a photon at particular energy: for example Co 60 emits at 1.173 MeV and 1.333 MeV. When Genie2k looks for these peaks at their known energy and it finds a real peak at 1.333 MeV, but not for 1.173 MeV because background is too high. However, for both, Genie2k attributes an efficiency and a error. We plotted these efficiencies and errors and compared it to the manual fits in figure 15. Genie2k's efficiency and error overlap the error fit only at energy where no peak was found (like 1.173 MeV in our example). When a peak is found, Genie2k uses another way to calculate the error, probably a fit of the error points. As the significant signal is given with peaks, the standard deviation of not found peaks is not of interest. Therefore, without knowing the use of error parameters, we concluded that it gives information on not found peak efficiency error. These parameters are thus not relevant and are not to be added to the database.

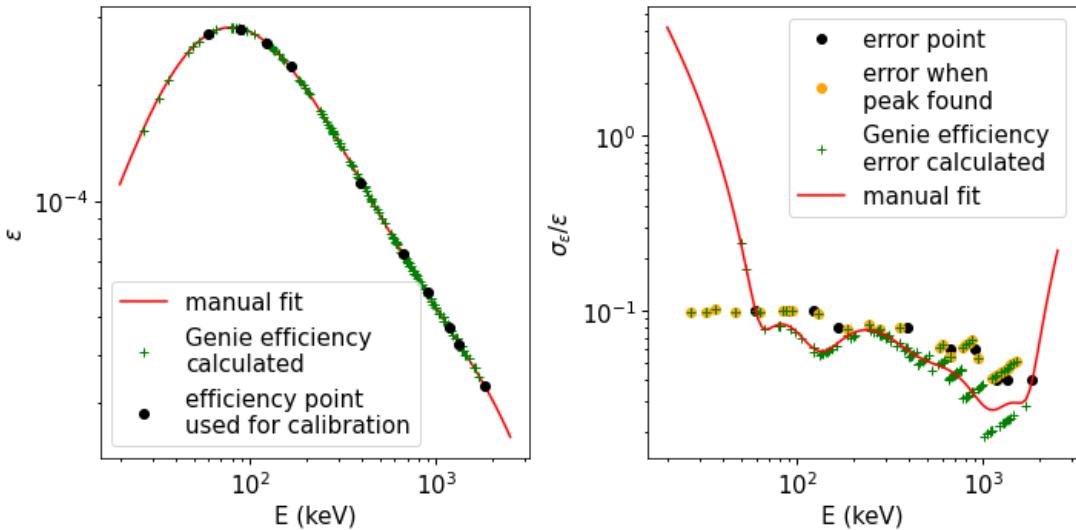


Figure 15: Comparison between manual fit and Genie2k attributed efficiency (left) and standard deviation (right). Manual fit overlap for not found peak signal, otherwise another method is used. Standard deviation parameter is thus not of interest, as it does not apply for significant signal.

### 3.2 Database and its implementing tool modification

Once the inventory was completed, the list of new parameters could be implemented in the CaTo1800 script and in the database structure. Most of the new data are general data about a spectrum, and thus were added to the current 'Spectrum' table. The error parameters (FWHM for each peak, activity for each nuclide) could also be added to the existing tables.

The calibration parameters (coefficient fit for energy, FWHM and efficiency) could be implemented in the current 'Spectrum' table, but this would have lead to a lot of empty column (the order of polynomial fit is not fixed for efficiency calibration). A choice was made to gather the calibration parameters into a new table named 'Calibration', with the foreign 'SpectrumID' key. The advantage of this structure is the possibility of extracting the order of the coefficient easily, and to avoid too many empty values. The columns are the key, the value of the coefficients, its order, its maximum order (order of the polynomial for the efficiency) and the type of calibration (energy, FWHM, logarithmic efficiency, and their corresponding errors). The new database structure is shown in figure 16.

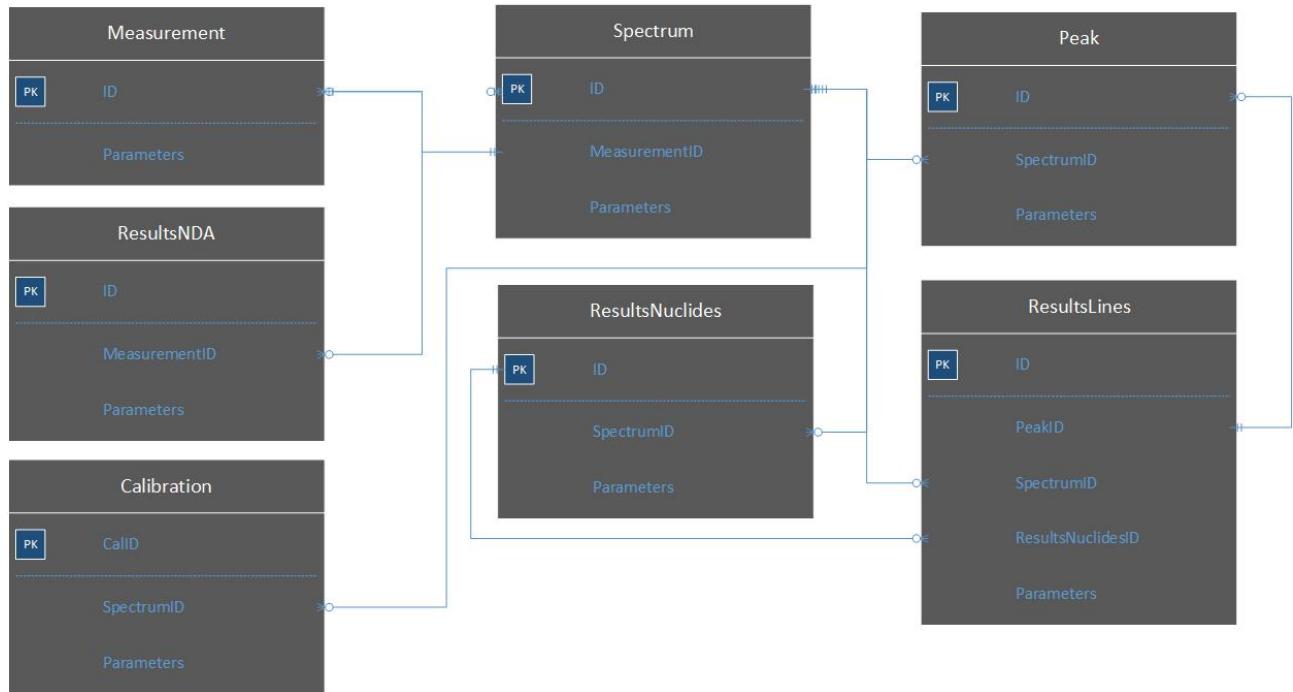


Figure 16: New database structure, with the new Calibration table linked with the SpectrumID. 80 parameters were added to the previous 90 ones.

Because the current database already contained data, the new database structure was implemented using SQL request, and for every row that were in the database, a "null" value was assigned for new parameters. Once the tables were added, the database was ready for receiving new data.

The next step consists in extracting these new parameters from the CNF files using the CaTo1800 tool. As shown previously, the data are stored in a class object. To acquire the new parameters, one should add a private and public property for each of them, and add them to the reading process. These script modifications were performed using a python 3 program that was adding code lines to the C# script. Finally, the stored procedures (SQL request applied by the C# script on the database) were updated.

CaTo1800 is now able to read and extract more parameters, and the database was configured to accept them. A guide was also written for the team giving the procedure to add a parameter to the data management (database and CaTo1800). This guide is shown in the appendix A.

### 3.3 User interface upgrade

The two previous subsections explained how new data were introduced in the data management. This subsection describes the need and the set up of new features for the data management.

CaTo1800 was first designed for reading and uploading CNF files in a database. Several limitations were underlined at the beginning of the internship:

- One by one process: CaTo1800 could read CNF files one by one, which was leading to a difficult implementation of the historic files (spectra analysed before the creation of the database) in the new database.
- No reading from the database: CaTo1800 was only reading files, so when a file was implemented, it was impossible to access these data with CaTo1800.

To enhance CaTo1800 operations, two virtual databases were created inside the application: one that contains some information of the read files, and another one containing some data of spectra in the database. These virtual databases allows to:

- Read multiple files ;
- Read the database ;
- Compare data between these virtual databases; and
- Update the database by pushing data from the virtual to the actual database.

These new features provide much more flexibility with the data, modifying the one way traffic of the data in a dialog between CaTo1800 and the database. The evolution of the features is shown in figure 17.

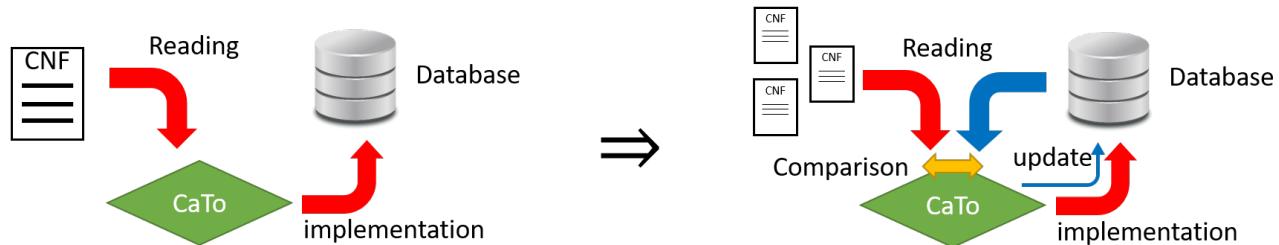


Figure 17: Previous and new CaTo1800 features: the application can now 1/ read multiple CNF files 2/ read from database 3/ compare data between CNF and database 4/ update database.

The new features introduce multiple application:

- Implementation of more data;
- Verification and traceability of the data
- Verification that the read files are not already uploaded in the database. Another comparison notifies whether a file with the same name is already in the database or not: this case can happen when a first analysis is performed and the file is uploaded. Afterwards another analysis of the same spectrum is made, which overwrites the data in the CNF file. The new file is thus a new analysis of a spectrum, and must be uploaded if needed. To distinguish between equal file in database and same name but not same analysis, we chose to compare the name and the analysis date (when both are equal, the read file is considered as already in the database). The analysis date is a new parameter added with this internship to the database;

- When a file is uploaded but the analysis or the measurement was wrongly performed, another file is uploaded, but the previous one should be marked as invalid, as no file can be deleted for traceability reasons. This "NotValid" property was previously modifiable in the database, and can now be modified from CaTo1800.

To implement these new features, a new user interface was designed and is shown in the figures 18 and 19. It is programmed using xaml language. For the main window, the upper list displays the read files, and the lower list shows the last uploaded files to the database. For the upload window, the upper list displays the read files, and the lower list shows the related files in the database (equal and same name but not equal). The colors refer to the status of the read files (green: not in the database; orange: same name in the database; red: equal in the database), this status is changed as green if the operator check the "NewVersion" property, otherwise he should check the "Avoid" property to prevent the red and orange files from the uploading and turn the color into gray. When a read file is selected, the related files in the database (displayed lower) are colored in lightblue. The operator can thus update the "NotValid" property and the comment.

All the information regarding these new features were written in a User guide, shown in the appendix B.

The screenshot shows the main window of the CaTo 1810 application. The title bar says "CaTo 1810". Below it is a menu bar with "File", "Upload", and "Help". The main area contains two tables. The top table has columns: "FileName", "AnalysisDate", "Geometry model", "NotValid", and "UpLoadDate". It contains three rows with different colors: red, orange, and green. The bottom table also has columns: "FileName", "AnalysisDate", "Geometry model", "NotValid", and "UpLoadDate". It contains several rows with various file names and dates.

FileName	AnalysisDate	Geometry model	NotValid	UpLoadDate
112-room-K1-01_01_20220301.CNF	2022-03-08 13:29:45	2022_010	<input type="checkbox"/>	Not upload
114-02-K2-01_01_20211130.CNF	2021-12-01 9:02:42	2021_116	<input type="checkbox"/>	Not upload
114-02-K3-01_01_20211130.CNF	2021-12-01 14:48:18	2021_116	<input type="checkbox"/>	Not upload

FileName	AnalysisDate	Geometry model	NotValid	UpLoadDate
BIS-K145_A_20220531.CNF	6/21/2022 9:38:10 AM	2021_154	<input type="checkbox"/>	6/21/2022 9:39:59 AM
114-00-K1-02_01_20211207.CNF	12/13/2021 9:51:04 AM	2021_227	<input type="checkbox"/>	6/21/2022 9:16:59 AM
114-05-K1-03_01_20211125.CNF	11/25/2021 9:04:25 AM	2020_373	<input type="checkbox"/>	6/20/2022 9:31:22 AM
114-01-K1-01_01_20211130.CNF	12/13/2021 9:52:09 AM	2021_116	<input type="checkbox"/>	6/20/2022 9:31:21 AM
114-00-K1-01_02_20211124.CNF	12/13/2021 9:50:14 AM	2021_129	<input type="checkbox"/>	6/20/2022 9:31:21 AM
114-00-K1-01_01_20211124.CNF	12/13/2021 9:49:14 AM	2021_129	<input type="checkbox"/>	6/20/2022 9:31:20 AM
112-room-K1-02_01_20220307.CNF	3/8/2022 1:30:04 PM	2022_010	<input type="checkbox"/>	6/20/2022 9:31:18 AM
114-00-K1-01_01_20211124.CNF	12/13/2021 9:49:14 AM	2021_129	<input type="checkbox"/>	6/20/2022 9:06:38 AM
112-room-K1-01_01_20220301.CNF	3/8/2022 1:30:45 PM	2022_010	<input type="checkbox"/>	6/16/2022 9:55:16 PM

Reading complete!

Figure 19: Upload window of the application, the above list is referring to the read files, the lower list to the related files in the database. The color indicates the status of the read files (green: not already in the database; orange: same same already in the database; red: same name and same analysis date in the database; gray: avoided from uploading). Select a read file colors the related files in light blue. The operator can modify the "NotValid" and Comment property of the read and the already in the database files.

The screenshot shows the upload window of the CaTo 1810 application. The title bar says "Upload data". Below it are two tables. The top table has columns: "FileName", "AnalysisDate", "Geometry model", "NotValid", "Avoid", "NewVersion", and "Comments". It contains three rows with different colors: red, orange, and green. The bottom table has columns: "FileName", "AnalysisDate", "Geometry mode", "NotValid", "UpLoadDate", and "Comments". It contains several rows with various file names and dates.

FileName	AnalysisDate	Geometry model	NotValid	Avoid	NewVersion	Comments
112-room-K1-01_01_20220301.CNF	2022-03-08 13:29:45	2022_010	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	
114-02-K2-01_01_20211130.CNF	2021-12-01 9:02:42	2021_116	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	
114-02-K3-01_01_20211130.CNF	2021-12-01 14:48:18	2021_116	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	

FileName	AnalysisDate	Geometry mode	NotValid	UpLoadDate	Comments
112-room-K1-01_01_20220301.C	3/8/2022 12:00:00 AM	2022_010	<input type="checkbox"/>	6/16/2022 12:00:00 A	
112-room-K1-01_01_20220301.C	3/8/2022 12:00:00 AM	2022_010	<input type="checkbox"/>	6/16/2022 12:00:00 A	
112-room-K1-01_01_20220301.C	3/8/2022 12:00:00 AM	2022_010	<input type="checkbox"/>	6/16/2022 12:00:00 A	
112-room-K1-01_01_20220301.C	3/8/2022 1:29:45 PM	2022_010	<input type="checkbox"/>	6/16/2022 2:18:25 PM	
114-02-K2-01_01_20211130.CNF	12/1/2021 12:00:00 A	2021_116	<input type="checkbox"/>	6/14/2022 12:00:00 A	

Reset Upload to DB

Figure 18: Main window of the application, the above list is referring to the read files, the lower list to the last uploaded files. The color indicates the status of the read files (green: not already in the database; orange: same same already in the database; red: same name and same analysis date in the database).

## Conclusion

This internship aimed to address the data management upgrading for radiological characterisation. Two main strategies have been performed: add more data to the process, by reading and implementing more parameters from the analysis ; and by increasing the features of the previous CaTo1800 tools. A range of 80 parameters were added to the process, which collects much more data about calibration for measurement and analysis reports. The new features implemented to CaTo1800 increases the efficiency, the accuracy and the traceability of the activity results.

However, the outcomes of this work does not affect only the workflow: during the internship, a large range of tools and material was produced to allow further development of the process. An excel sheet giving all the parameters extractable from data file, some python scripts to update easily the process and add new parameters, some report explaining current features, updating procedures and further development strategies. This internship gave thus an overview of the possibilities regarding data management and provided some tools to easily update the CaTo1800 tool.

The perspectives for data management in the long run are numerous. The addition of new parameters underlined the need of a better and wider understanding of the parameters available in the encrypted data files. The new features applied to the process gave also some ideas to increase the flexibility of data manipulation. Adding more comparison tools, implement more data (identification ones such as material name etc.) would centralise all the data processing around this new application. This goal recalls the LiNDA project at the SCK CEN. This project aims at gathering all the data from nuclear material measured with laboratory set up in one big system. This project centralises all the measurement process, analysis calculation, data management and regulation compliance report conception in one comprehensive application. This project provided the first CaTo1800 program, and an integration of the new CaTo1800 in this project would be beneficial for the characterisation team in BR3 as for the LINDA project, that would enhance its features to in situ measurement.

## References

- [1] Status and Trends in Spent Fuel and Radioactive Waste Management. Number NW-T-1.14 (Rev. 1) in Nuclear Energy Series. INTERNATIONAL ATOMIC ENERGY AGENCY, Vienna, 2022.
- [2] Gordon Gilmore. Practical gamma-ray spectroscopy. John Wiley & Sons, 2008.
- [3] Renée Cools. Assessment of the radiological characterization strategy during decommissioning based on preliminary data analysis. Master's thesis, 2021.
- [4] John Neter William Li Michael H. Kutner, Christopher J. Nachtsheim. Applied Linear Statistical Models. The McGraw-Hill Companies, 2004.

# A Update procedure

## CaTo modification guide

SERANDOUR Antoine

This report aims to give all the steps to implement and update the software CaTo and the ISOCSBR3 database. The first section deals with the implementation of CaTo on a new database, and the second section refers to the addition of a new parameter to the process.

### 1 Implement CaTo

This section highlights the key steps for implementing CaTo to a new database. The implementation is divided in three times : modification of the database structure, creation of the stored procedures and connecting CaTo to the new database.

#### 1.1 Database modification

The structure of the database should be the same as the current structure of the working database of CaTo. The tables and the columns can be created using the following SQL query.

Code 1: Table creation
<code>CREATE TABLE TableName (     Column1 DataType1,     Column2 DataType2,     Column3 DataType3,     Column4 DataType4 )</code>

Code 2: Column creation
<code>ALTER TABLE TableName ADD ColumnName DataType;</code>

Code 3: Column data type modification
<code>ALTER TABLE TableName ALTER COLUMN ColumnName DataType</code>

Note that there is no query for changing the name of a column. However, SQL server displays a user interface to modify the names, if the accessibility is sufficient. Another way to modify a column name is to delete and create a new column, as for modifying a table name.

Code 4: Table suppression
<code>DROP TABLE TableName;</code>

Code 5: Column suppression
<code>ALTER TABLE TableName DROP COLUMN ColumnName;</code>

Once the database structure is the same as the previous one, the stored procedures can be added to the new database.

#### 1.2 Stored Procedures implementation

The stored procedures are SQL queries that a code can apply on the database. They are thus written in the initial database (in programmability/stored procedures for SQL server). If the names of the columns and tables of the new database are the same as in the initial one, the stored procedure can be

copied. With a right click on the initial stored procedure on SQL server, the modify feature appears and is selected. A query appears and should just be slightly modified as follow :

Code 6: Stored procedure modification

```
USE [InitialDatabase]
GO
/****** Object: StoredProcedure
[dbo].[uspName]
Script Date: uspDate *****/
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON
GO
-----
-- Author: <Author,,Name>
-- Create date: <Create Date,,>
-- Description: <Description,,>
-----
ALTER PROCEDURE [dbo].[uspName]
@uspParam1 DataType1,
@uspParam2 DataType2
AS
BEGIN
    ...
END
```

Code 7: Stored procedure creation

```
USE [NewDatabase]
GO
/****** Object: StoredProcedure
[dbo].[uspName]
Script Date: uspDate *****/
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON
GO
-----
-- Author: <Author,,Name>
-- Create date: <Create Date,,>
-- Description: <Description,,>
-----
CREATE PROCEDURE [dbo].[uspName]
@uspParam1 DataType1,
@uspParam2 DataType2
AS
BEGIN
    ...
END
```

The stored procedures to add are :

- Insert stored procedures (as many as tables in the database)
- usp.PrintDatabase
- usp.TestOtherFileInDB

### 1.3 CaTo connection

Finally, the initial stored procedures should be replaced in CaTo by the new one, using the xsd file.

For the insert stored procedures, they can be added to the Query table adapter using the "Add Query" button appearing with a right click. For each, the connection should be specified, linking CaTo to the new database. A new window will appear asking for the command type to use, that should be "use existing stored procedures". Afterwards the stored procedures is selected. The Name of the function should be the same as the one used for the initial stored procedure. Finally, the initial procedure can be deleted.

For the PrintData and TestOtherFile procedures, a new table adapter can be created using right click on an empty region. First the connection is specified, then the command type "use existing stored procedures" should be checked and finally the "select" blank has to select the PrintData or TestOtherFile procedure. The initial datatable can be removed. The only table that is not modified is the DataTableCandidate.

Finally, some modification of the C# code has to be performed. In both the MainWindow.cs and Window1.cs codes, the mention of the initial database must be replaced by the new database. Afterwards, CaTo should run without any errors.

## 2 Add a parameter to CaTo

To add a new parameter to the extraction process, the database, the sorted procedures and the caTo code should be modified.

### 2.1 Database and procedures modification

First, a new column should be added to a table. This can be done using the SQL queries written in the 1.1.

Afterwards, the stored procedure inserting data to the table should be modified. To do so, a right click on the stored procedure (programmability/stored procedure) allow its modification. The new parameter can be added to the initial procedure as shown in the following :

Code 8: Stored procedure modification

```
ALTER PROCEDURE [dbo].[uspName]
@uspParam1 DataType1,
@uspParam2 DataType2
AS
BEGIN
    INSERT INTO TableName
    (uspParam1 ,uspParam2)
    VALUES
    (@uspParam1 ,@uspParam2)
END
```

Code 9: Stored procedure after modification

```
ALTER PROCEDURE [dbo].[uspName]
@uspParam1 DataType1 ,
@uspParam2 DataType2,
@uspParam3 DataType3
AS
BEGIN
    INSERT INTO TableName
    (uspParam1 ,uspParam2,uspParam3)
    VALUES
    (@uspParam1 ,@uspParam2,
    @uspParam3)
END
```

### 2.2 Modification of the C# code

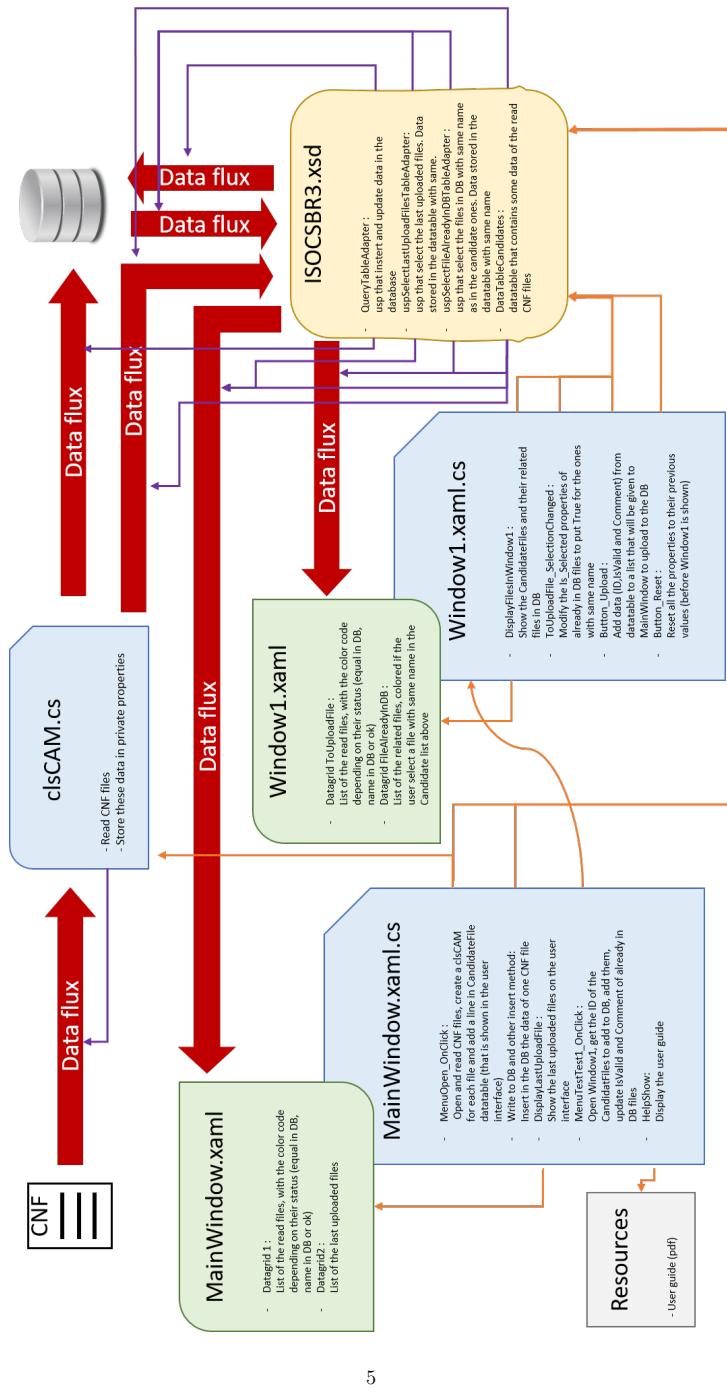
Finally, the CaTo code should be modified to extract and upload the new parameter.

The first step is to update the stored procedures of CaTo, using the xsd file (click then right click on the new stored procedure and select "Configure" and "Finish").

Afterwards, the C# code should implement the new parameter. this can be done following the next steps:

1. Create a private property (*clsCAM*) :
  - Copy the current structure line beginning with *private*
  - Change the CAM request
  - Change the type with respect to the type returned by the CAM request
2. Create a public property (*clsCAM*) :
  - Copy the current structure lines beginning with *public*
  - Change the name of the property with respect to the name of the CAM request
  - Change the m\_CAM request for the *get* and *set* methods
  - Change the type with respect to the type returned by the CAM request
3. Add the data into a reading method for common parameter (*clsCAM*) :

- Copy the current structure lines beginning with *try*
  - Change the name of the property with respect to the name of the CAM request
  - Change the m\_CAM request for the *get\_Param* methods
  - Change the type with respect to the type returned by the CAM request
4. Add the data into a reading method for record parameter (*clsCAM*) :
- Copy the current structure lines beginning within the *for* loop
  - Change the name of the property with respect to the name of the CAM request
  - Change the m\_CAM request for the *get\_Param* methods
  - Change the type with respect to the type returned by the CAM request
  - Create a new line before defining a new list with the template  
"m\_CAM\_X\_XXXX = new List<XXX>();"
5. Modify the input of the stored procedures (*MainWindow*) :
- Add the new parameter to the list of the input parameters, using the *clsCAM* class



## B User guide

### CaTo user guide

CaTo is an application that reads CNF files to send interesting data to the ISOCSBR3 database. It is based on the work of Wim De Boeck and was adapted to ISOCS measurement by Antoine Sérandour. This guide gathers its features.

#### Main Window

When opening CaTo, two data lists are displayed on the interface. The upper one will contain the candidate files to upload to the database, and the lower one contains the last 10 files added to the database.

To implement files to the candidate list, click on File and open. Several files can be selected. After reading, they will appear in the upper data list. They appear with a color depending on their status.

- green : the file is not in the database
- orange : the name of the file is already in the database
- red : the file seems to be already in the database

A file is considered already in the database if there is a spectrum in the database with the same spectrum table.

Candidate files can be sent to the database with clicking on upload then send to ISOCSBR3. A new window appears to give more details on the uploading.

The screenshot shows the CaTo 1810 application window. At the top, there is a menu bar with 'File', 'Upload', and 'Help'. Below the menu is a toolbar with several icons. The main area consists of two tables. The upper table has columns: 'FileName', 'AnalysisDate', 'Geometry model', 'NotValid' (checkbox), and 'UpLoadDate'. It contains three rows: 112-room-K1-01\_01\_20220301.CNF (red), 114-02-K2-01\_01\_20211130.CNF (orange), and 114-02-K3-01\_01\_20211130.CNF (green). The lower table has the same columns and contains a list of files, many of which are colored blue, indicating they are related to the candidate files in the upper list. At the bottom of the window, a message says 'Reading complete !'.

FileName	AnalysisDate	Geometry model	NotValid	UpLoadDate
112-room-K1-01_01_20220301.CNF	2022-03-08 13:29:45	2022_010	<input type="checkbox"/>	Not upload
114-02-K2-01_01_20211130.CNF	2021-12-01 9:02:42	2021_116	<input type="checkbox"/>	Not upload
114-02-K3-01_01_20211130.CNF	2021-12-01 9:48:18	2021_116	<input type="checkbox"/>	Not upload

FileName	AnalysisDate	Geometry model	NotValid	UpLoadDate
BIS-K145_A_20220531.CNF	6/21/2022 9:38:10 AM	2021_154	<input type="checkbox"/>	6/21/2022 9:39:59 AM
114-00-K1-02_01_20211207.CNF	12/13/2021 9:51:04 AM	2021_227	<input type="checkbox"/>	6/21/2022 9:16:59 AM
114-05-K1-03_01_20211125.CNF	11/25/2021 9:04:25 AM	2020_373	<input type="checkbox"/>	6/20/2022 9:31:22 AM
114-01-K1-01_01_20211130.CNF	12/13/2021 9:52:09 AM	2021_116	<input type="checkbox"/>	6/20/2022 9:31:21 AM
114-00-K1-01_02_20211124.CNF	12/13/2021 9:50:14 AM	2021_129	<input type="checkbox"/>	6/20/2022 9:31:21 AM
114-00-K1-01_01_20211124.CNF	12/13/2021 9:49:14 AM	2021_129	<input type="checkbox"/>	6/20/2022 9:31:20 AM
112-room-K1-02_01_20220307.CNF	3/9/2022 1:30:04 PM	2022_010	<input type="checkbox"/>	6/20/2022 9:31:18 AM
114-00-K1-01_01_20211124.CNF	12/13/2021 9:49:14 AM	2021_129	<input type="checkbox"/>	6/20/2022 9:06:38 AM
114-00-K1-01_01_20220301.CNF	3/9/2022 1:30:45 PM	2022_010	<input type="checkbox"/>	6/20/2022 9:45:46 AM

#### Upload data window

This window is composed of two data lists : the upper list gathers the candidates files. The lower list displays the files already in the database that are related to the candidates files. When selecting a candidate file, the related files (same file name) in the database are colored in blue.

The operator can select the uploading properties ("Avoid" and "NewVersion" checkbox) for each read files.

- Check "Avoid" : the file will be excluded from uploading to the database. The file will appear in gray.

- Check "NewVersion" : for red or orange files, this property confirm the uploading of these files.  
The file will be displayed in green

The properties "Avoid" and "NewVersion" cannot be both checked, otherwise an error message will be displayed while uploading to database.

The operator can add some comment to the files to display in the database. It is also possible to modify the "Notvalid" property, either of candidate files or related files already in the database.

Let's consider several examples :

- A file is already in the database with the same name but was analysed with another geometry model. It will be shown in orange, and the related file already in the database is light blue. The operator can click on "Newversion" to force the upload, and can mark the light blue file as "NotValid".
- A file is already in the database twice, with the same name and analysis date. It will be colored in red, and two files in the lower datalist are light blue. The operator clicks on "Avoid" and thus prevent the uploading of this file, and click as "NotValid" one of the two files already in the database.
- A candidate file is green. The operator can comment it and directly click on upload, without clicking on "NewVersion".

After setting the properties for uploading, the operator can click on "Upload to database" button. A "Reset" button reset all the modifications of "Avoid", "NewVersion", "NotValid" and "Comment".

The screenshot shows a software interface titled "Upload data". At the top, there is a toolbar with icons for back, forward, and search. Below the toolbar are two tables of data:

FileName	AnalysisDate	Geometry mode	NotValid	Avoid	NewVersion	Comments
112-room-K1-01_01_20220301.CN	2022-03-08 13:29:45	2022_010	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
114-02-K2-01_01_20211130.CNF	2021-12-01 9:02:42	2021_116	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	
114-02-K3-01_01_20211130.CNF	2021-12-01 14:48:18	2021_116	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	

FileName	AnalysisDate	Geometry mode	NotValid	UploadDate	Comments
112-room-K1-01_01_20220301.C	3/8/2022 12:00:00 AN	2022_010	<input type="checkbox"/>	6/16/2022 12:00:00 A	
112-room-K1-01_01_20220301.C	3/8/2022 12:00:00 AN	2022_010	<input type="checkbox"/>	6/16/2022 12:00:00 A	
112-room-K1-01_01_20220301.C	3/8/2022 12:00:00 AN	2022_010	<input type="checkbox"/>	6/16/2022 12:00:00 A	
112-room-K1-01_01_20220301.C	3/8/2022 12:09:45 PM	2022_010	<input type="checkbox"/>	6/16/2022 2:18:25 PN	
114-02-K2-01_01_20211130.CNF	12/1/2021 12:00:00 A	2021_116	<input type="checkbox"/>	6/14/2022 12:00:00 A	

At the bottom of the interface, there are two buttons: "Reset" and "Upload to DB".

Structured Query Language (SQL)  
Pressurized Water Reactor (PWR)  
High Purify Germanium (HPGe)  
In Situ Object Calibration Software (ISOCS)  
Full Width at Half Maximum (FWHM)  
Conjunctive Normal Form (CNF)  
Configuration Access Method  
(CAM)  
Mean Square Error (MSE)  
Least Common Multiple (LiNDA)  
Belgian Reactor 3 (BR3)  
Studiecentrum voor Kernenergie  
Centre d'Etude de l'Energie Nucléaire (SCK CEN)