SAND REPORT

SAND2007-1841
Unlimited Release
Printed March 2007

Capabilities, Methodologies, and Use of the Cambio File-Translation Application

Dr. George P. Lasche

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

Sandia is a multiprogram laboratory operated by Sandia Corporation,
a Lockheed Martin Company, for the United States Department of Energy's
National Nuclear Security Administration under Contract DE-AC04-94AL85000.

Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.

Available to DOE and DOE contractors from
U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone: (865) 576-8401
Facsimile: (865) 576-5728
E-Mail: reports@adonis.osti.gov
Online ordering:http://www.doe.gov/bridge

Available to the public from
U.S. Department of Commerce
National Technical Information Service
5285 Port Royal Rd
Springfield, VA 22161

Telephone: (800) 553-6847
Facsimile: (703) 605-6900
E-Mail: orders@ntis.fedworld.gov
Online order:http://www.ntis.gov/help/ordermethods.asp?loc=7-4-0#online

SAND2007-1841
Unlimited Release
Printed March 2007

Capabilities, Methodologies, and Use of the Cambio File-Translation Application

Dr. George P. Lasche

Sandia National Laboratories
P.O. Box 5800
Albuquerque, NM 87185

Abstract

This report describes the capabilities, methodologies, and uses of the Cambio computer application, designed to automatically read and display nuclear spectral data files of any known format in the world and to convert spectral data to one of several commonly used analysis formats. To further assist responders, Cambio incorporates an analysis method based on non-linear fitting techniques found in open literature and implemented in openly published source code in the late 1980s. A brief description is provided of how Cambio works, of what basic formats it can currently read, and how it can be used. Cambio was developed at Sandia National Laboratories and is provided as a free service to assist nuclear emergency response analysts anywhere in the world in the fight against nuclear terrorism.

Table of Contents

Figure

Introduction

Cambio was developed in response to a need of nuclear emergency response analysts everywhere to be able to quickly read the data from any of a widely growing number of formats used by both commercial and government detector systems worldwide. As the number of manufacturers of nuclear detection instrumentation grows, so does the number of data formats that must be able to be read by emergency analysts. Manufacturers of instrumentation often need to create new and more complex versions of their own formats as technology advances and as new user requirements lead to new, more sophisticated instruments. Cambio is a stand-alone executable written in Borland Delphi® Object Pascal, runs under Microsoft Windows® XP, and requires no special installation or any supporting files. New compilations of Cambio are frequently made available as the ability to read these new formats are added to Cambio.

The file translation capability of Cambio

How the file translation capability of Cambio works

Cambio has code that extracts basic elements of information from roughly 52 known spectral formats and their variants. Files are first tested to see if they have nearly 100% ASCII text content. Tests then branch into two main groups: those in ASCII and those that are binary. The following tests are conducted for both ASCII and binary files:

- **Determine file format:** Some formats do not have any designated file extension, some file extensions are changed in saving to disk or in transmission, and some file extensions are used for several very different formats. For these reasons, Cambio first reads a portion of each file for specific signatures of its format without regard to file extension, and tries the format it thinks most likely.

- **Test identified format:** When the format is read, the result is subjected to a series of tests appropriate for that format. These include tests for reasonableness of number of channels, energy calibration, live time, real time, and acquisition date and time. Depending on the format, failure to pass these tests results in either substitution of the quantity by a default value (with warning), or with rejection of consideration for that particular format (if the failure is considered severe).

- **Determine format of rejected files:** If the format under trial is rejected, Cambio then tries the next most likely formats based on the file extension. If this fails, Cambio tries every format, one at a time, beginning with the most common formats, until one of them appears to work. If this fails, Cambio considers the format as unknown and attempts to extract and display the raw spectral data.

- **Extract and display data of files with unknown format:** ASCII files of unknown format are searched column-wise and row-wise for a series of ASCII numbers that meet a set of criteria to qualify as spectra. In the same way, if a binary format can not be recognized by this process, Cambio will nevertheless attempt to extract and display the raw spectral data from each spectrum contained in the file, even if the file is of an unknown binary type. This is done by searching for the largest contiguous sequence

of non-negative numbers beginning with 4-byte integer types, followed by two-byte integer types, followed by four-byte floating point types. The entire set must then pass a chi-square test to see if it fits the Poisson nature of a nuclear spectrum. If that test is passed, that portion of the sequence having members numbering the largest integer power of two and having a reduced chi-square most approximating unity is selected and displayed. The user is then presented with a display of the data and an option to continue the search for another possible spectrum within the file.

What information is extracted?

If found, the following data are displayed:

- Raw spectral data
    - If an energy calibration is known, the data are displayed as counts/keV.
    - The vertical scale may be in either log-base-10 of counts, which is the default, square root of counts, or counts
    - The horizontal scale may be in either keV or channels
- Number of channels of data
- Date and time of acquisition
- Live time in seconds
- Real time in seconds
- Comments (two separate comment lines are possible)
- The record number of the current spectrum and total number of spectral records in the file (because some formats contain more than one spectral record)
- Total counts and counts per second of live time in the current view
- Energy calibration parameters up to the quadratic term

Formats currently recognized by Cambio

Cambio currently recognizes the following formats:

- ANSI N42.42 (only the essential data customarily read and saved by Cambio are implemented)
- Aspect MKC
    - A-02
    - A-03
- Berkeley Nucleonics SAM-935 (*.ans) with QCC decompression and linearization
- Berthold LB-125 (includes multi-record files)
- Bubbletech
- Canberra
    - Generic CAM format (*.cnf) (includes detection of “counterfeit” variants)
    - Accuspec (*.dat)
    - Toolkit (*.tka)
    - Inspector 1000 specific CAM format (*.cnf)

- CTC “MCS”
- Davidson (4 variants)
- Exploranium (includes multi-record files)
    - Gr-130
    - Gr-135 v1 (2 variants)
    - Gr-135 v2 (2 variants)
    - ASCII (3 variants)
- FieldSpec (now IdentiFinder) (includes the new FieldSpec-N variant)
    - Native format (*.spc)
    - IAEA SPE variant (*.spe)
- GADRAS (includes multi-record files)
    - PCF
    - PCC (RAID, RIS, SMART)
    - ASC
- IAEA Generic SPE
    - 3 methods in common use of representing date and time
    - 3 methods in common use of representing polynomial energy calibration and the method of inferring energy calibration from a series of channel-energy pairs
    - Mini-MCA (*.spe)
- LANL
    - GN-2 (garble-resistant; multirecord)
    - GN-3 (4 variants) (resistant to midstream restarts, multi-record)
    - Palm Pilot
- ORTEC
    - CHN
    - SPC
    - SPE
    - Print-to-file ASCII
- PDR-78 (file extensions are actually sequence numbers)
- PGT Avalon (*.ans)
    - Revision 1
    - Revision 2 and above
- Quantrad Ranger (multi-record)
- Rainbow Model 7010
- RobFit (FREE, HDTA, Z4DA; real and integer)
- SAIC RadSmart
- STE Pager-X
- STL

    - Cadillac (ASC and CSV)
    - Yugo
- Target NanoSpec
- XIA Polaris (*.itx)
- XRF ICS-4000 (2 variants)

Graphic views in Cambio

Users can enlarge and reduce the view to examine data at higher or lower detail as follows:

- To enlarge the spectral region to the full view ("zoom"):
    - Use the left mouse button to drag an outline of the desired area starting from the upper left and ending on the lower right.
    - Release the mouse button to allow the area to fills the full view.
    - Repeat to enlarge the view again as often as desired.
- To return to the original full view
    - Use the left mouse button to drag an outline anywhere in the display from the lower-right to the upper-left.
    - Release the mouse button to restore the original view.
- To scroll the expanded view to see adjacent data ("pan")
    - Use the right mouse button anywhere in the display to drag the view window in the desired direction.

Tips for efficient use

The following tips can help optimize user efficiency:

- To open files easily:
    - Place a shortcut to Cambio on the Windows desktop, then drag any spectrum file onto the icon and Cambio will automatically open the spectrum file.
    - Alternatively, use a Cambio menu item in the File menu to associate file extensions of common spectral formats with Cambio. Then, open the files with Cambio by double-click on a spectral file.
- To page to a particular spectral record in a large multi-record file
    - Type in the record number directly and then use the Enter key to go directly to the desired record.
- To change any color:
    - Drag the desired color in the Display tab page to the name of the graphic element below the color palette.
    - Return to the defaults from the Display menu.
- To calibrate a spectrum
    - Zoom first in the region of a good peak at high energy, preferably the 2614-keV peak of Tl-208.
    - On the Calibration tab page, click the Fit Peak button, then click the Apply button, then unzoom.
    - Repeat the process working toward lower energies with as many peaks as desired.

- Alternatively, apply a “typical” calibration from any of several known instruments from the Calibration menu.

- To compare two spectra or to use Spectrum Operations like time-scaled subtraction:
  - Make the active spectrum a reference spectrum with the “Exchange the Primary Spectrum with the Reference Spectrum” menu item in the Spectrum Operations menu.

How the optional analysis capability of Cambio works

(Note: The material in this section was the subject of a talk presented to the INMM on May 12, 2005, and was approved as Unclassified for Unlimited Release by the Sandia National Laboratory Review and Approval process, Document Number 5232124, SAND Number 2005-2924 C**. The following information is extracted from that SAND report.)**

Non-Linear Fitting of Nuclear Spectra

The non-linear method of fitting nuclear spectra in Cambio is best described as whole spectrum non-linear least-squares fitting with advances made by Dr. Bob Coldwell from 1982 to 1991. Dr Coldwell is an Associate Professor of Physics at the University of Florida.

The method is quite powerful as can be seen from the deconvolution of interfering peaks in this sample analysis of a spectrum (this graphic was produced by an actual Cambio analysis):

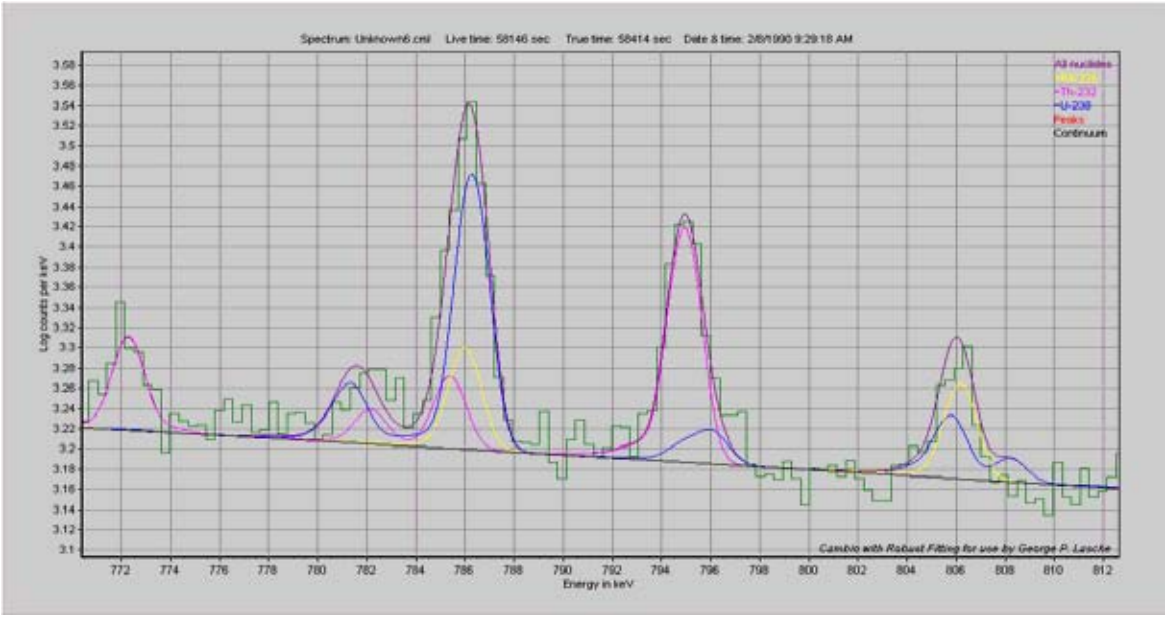

Figure 1. An example of the power of non-linear least squares whole-spectrum fitting to resolve interfering peaks.

Chi-Square as a Measure of Goodness of Fit

In the Cambio method of analysis, chi-square is used as a measure of goodness of fit. This quantity is the sum of the squares of the differences between the fit at the center of each channel of data and the value of each channel of data in the fitted range, less the number of free parameters to be fitted (known in statistical analysis as "degrees of freedom." By setting the derivative of chi-square with respect to each free parameter to zero, a set of simultaneous equations is generated for solution. If all of the coefficients in the resulting matrix are independent of all the other parameters, a linear matrix will result with a straightforward trivial solution. Chi-square is mathematically represented as follows:

$$\chi^2 = \sum_i \left(\frac{f_A(\vec{c}, x_i) - f_i}{\varepsilon_i} \right)^2$$

where f_A is the fitting function, the vector c represents the coefficients of the free parameters, x is the channel number, and f_i is the number of counts in that channel.

The Non-Linear Fitting Problem

However, if in addition to the activities of the component radionuclides, the energy calibration, detector efficiency response, resolution function, or attenuation by absorbing materials are to be varied, then all parameters have a dependency of the values of the others and there results a non-linear problem that must be solved in iterative steps. The challenge is to arrive at the global minimum in the chi-square hyperspace, and to do so quickly enough to be of practical use. Two approaches are used. First, the "method of steepest descent" that follows the negative of the gradient in linear steps. Then, when no further improvement is possible, the "exact solution" by linear matrix inversion under the assumption that the local minimum is near enough that all quantities vary at most quadratically as they deviate from the true minimum. Two problems typically exists with these approaches respectively: 1) if change is so slow near the minimum that it is within machine error (the "death valley" problem), and 2) how to treat the resulting term with the 2nd derivative.

The Selection of the Starting Point is Critical

It is essential to select a starting point near enough to the global minimum that it can be found with either technique. In Cambio this is done by manual selection and identification of at least two peaks. From this information, energy calibration and resolution calibration approximations are automatically made. The code, during analysis, will expand each non-linear function in a Taylor series about the initial starting vector as follows:

$$\left.\frac{\partial \chi^2}{\partial c_M}\right)_{\vec{c}} = \left.\frac{\partial \chi^2}{\partial c_M}\right)_{\vec{c}_0} + \sum_L \left(c_L - c_{L,0}\right) \left.\frac{\partial^2 \chi^2}{\partial c_M \partial c_L}\right)_{\vec{c}_0} + \quad \cdots$$

The Marquardt Parameter

Each of the diagonal elements of the second derivative of $\chi 2$ matrix to be inverted is multiplied by a factor $(1 + \lambda)$, where λ is a constant known as the Marquardt parameter.

Coldwell adds a variable multiplier S to account for the fact that the nonlinearities can be very different for each of the parameters C. The elements of S are later explicitly solved for to speed convergence.

When λ is large, the matrix is diagonally dominant, equivalent to forcing the derivatives of χ^2 with respect to each parameter to become independent, and the solution tends toward the method of steepest descent. When λ is small, an exact solution for the C is obtained provided we are in the range of quadratic convergence.

When this method may be of best use

This method may be of best use in any of the following cases:

- Detector characteristics not known
- No background available
- Desired information is sufficiently contained in the peaks, so the shape of the underlying continuum can be ignored.
- Peaks of interest masked by larger interfering peaks
- Particularly well-suited for HPGe spectra

Why not just find peaks?

In applications where peaks are not reliably identifiable, non-linear least-squares fitting can results in much higher confidence of detection. Examples include:

- Sparse data (too few counts)
- Noisy data (background radiation noise is high)

The rigorous mathematical uncertainty in the result of each parameter is readily found. The uncertainty in result for correlated quantities (such as the ratio of the activities of two radio-nuclides) can be determined with greatly reduced uncertainty by taking correlations into account.

Differences from Peak Fitting

Peaks are not searched for and then matched to a library. Instead –

- A single coefficient for the activity of each nuclide and a coefficient for each calibration, attenuation, and efficiency term are varied together until a best fit to all of the data is found.
- At each search step, a sum of whole-spectrum above-the-continuum nuclide shapes is generated and compared to the data.

Resources and References

Cambio file format translation program, available from:

George Lasche, gplasch@sandia.gov

Basics of non-linear least-squares fitting:

Bevington, P.R. 1969, Data Reduction and Error Analysis for the Physical Sciences, (New York: McGraw-Hill).

Press, W.H. et al., Numerical Recipes (Sec 15.5), Cambridge University Press.

Advances by Dr. Robert L. Coldwell (1982-1991)

Coldwell, R.L. and Bamford, G.J. 1991,The Theory and Operation of Spectral Analysis Using ROBFIT (New York: American Institute of Physics)

Bob Coldwell, ufbobc@ufl.edu

Source code in FORTRAN is at www.phys.ufl.edu/~coldwell

Distribution

1	MS1210	George Lasche, 5935
1	MS9004	William Ballard, 8100
1	MS9152	Noel Nachtigal, 8962
1	MS9004	Pat Falcone, 8110
2	MS9018	Central Technical Files, 8944
2	MS0899	Technical Library, 4536