

---

**Pacific Northwest  
National Laboratory**

Operated by Battelle for the  
U.S. Department of Energy

RECEIVED

MAR 31 1999

OSTI

**Existing Data Format for  
Two-Parameter Beta-Gamma  
Histograms for Radioxenon**

T.W. Bowyer  
T.R. Heimbigner  
J.I. McIntyre  
A.D. McKinnon  
P.L. Reeder  
E. Wittinger

February 1999

Prepared for the U.S. Department of Energy  
under Contract DE-AC06-76RLO 1830



## **DISCLAIMER**

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, make any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

## **DISCLAIMER**

**Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.**

# EXISTING DATA FORMAT FOR TWO-PARAMETER BETA-GAMMA HISTOGRAMS FOR RADIOXENON

T.W. Bowyer, T.R. Heimbigner, J.I. McIntyre,  
A.D. McKinnon, P.L. Reeder, E. Wittinger

Pacific Northwest National Laboratory

February 25, 1999

## Introduction

There is a need to establish a commonly acceptable format for storing beta-gated coincidence data for stations in the International Monitoring System (IMS) for the Comprehensive Nuclear-Test-Ban Treaty (CTBT). The current aerosol RMS type data format is not applicable for radioxenon in that the current format contains implicit assumptions specific to conventional gamma-ray spectrometry. Some assumptions in the current RMS format are not acceptable for the beta-gated spectra expected from the U.S. Department of Energy PNNL Automated Radioxenon Sampler-Analyzer (ARSA) and other similar systems under use or development from various countries.

The RMS data format is not generally applicable for radioxenon measurements in the CTBT for one or more of the following main reasons:

- 1) The RMS format does not currently support 2-dimensional data. That is, the RMS data format is setup for a simple 1-dimensional gamma-ray energy histogram. Current data available from the ARSA system and planned for other radioxenon monitors includes spectral information from gamma-rays and betas/conversion electrons. It is worth noting that the beta/conversion electron energy information will be used to separate the contributions from the different radioxenons.
- 2) The RMS data format assumes that the conversion between counts and activity can be calculated based (in part) on a simple calibration curve (detector efficiency curve) that depends only on energy of the gamma-ray. In the case of beta-gated gamma-ray spectra and for 2-dimensional spectra, there are generally two detector calibration curves that must be convoluted, the lower energy cutoff for the betas must be considered, and the energy acceptance window must be taken into account to convert counts into activity.
- 3) The RMS format has header information that contains aerosol-specific information that allows the activity (Bq) calculated to be converted into a concentration (Bq/SCM). This calculation is performed by dividing the activity calculated (Bq) into number of standard cubic meters of air (SCM) passed through the filters. Most xenon samplers do not have a 100% collection and transfer efficiency, and these efficiencies should not be assumed constant, so that the total volume flow through the sampler may not be used to convert activity into concentration. There is a pretty straightforward analogy that requires, instead, the total volume of xenon gas measured by the xenon station for the conversion.

The following paper describes one possible file format for storing the multi-parameter beta-gamma coincidence spectra generated by the DOE PNNL ARSA sampler. This format proposal was generated as a draft guide to begin discussions.

## Format Features:

The proposed format, "Xenon Monitoring Systems", XMS for short, contains the information necessary to convert a spectrum when combined with a background spectrum and other ancillary data contained in the data header into a concentration for the xenon radionuclides  $^{131m}\text{Xe}$ ,  $^{133g}\text{Xe}$ ,  $^{133m}\text{Xe}$ , and  $^{135}\text{Xe}$ . The general data format consists of a data header which includes information on detector efficiencies, dates and times of collections, and other ancillary information, and a 100 x 256 data array. Other format features include:

- ASCII data encoding. This eliminates byte ordering and structure alignment concerns common with raw binary files.
- Designed to support two-parameter beta/gamma histograms but can also be used for gamma-only histograms (in which case the number of beta channels is set to '1' indicating one large beta window.
- Free form, comma delimited data section is space efficient making the file not much larger than its raw binary equivalent.

The format is ASCII encoded and uses line-feed characters (ASCII(10)) to denote end-of-line. Blank lines and lines in which the first non-whitespace character is a pound character ('#') are ignored. The file consists of header and data sections as described below. The file size is expected to be on the order of 50K uncompressed, but as small as 3K when compressed.

## FORMAT DETAILS: HEADER SECTION

The header section contains a series of tagged fields with information about the file, the sample station, the detectors and the sample itself. A typical header might appear as follows:

```

FORMAT: XMS1.0
STATION: XNG01
GAMMA-CHANNELS: 256
GAMMA-GAIN/OFFSET: 1.000/23.000
BETA CHANNELS: 100
BETA-GAIN/OFFSET: 2.00/3.00
EFFICIENCY-WINDOWS:
131m:150.0/200.0/28.0/32.0/0.75
133m:210.0/230.0/28.0/32.0/0.85
133g1:3.000/250.0/28.0/32.0/0.59
133g2:3.000/400.0/78.0/84.0/0.95
135g:3.000/1000.0/240.0/260.0/0.85
STABLE-XENON-VALUE: 1.6
TYPE: SAMPLE
DATE: 1998/12/10 09:44:05 (GMT)
SAMPLE-START: 1998/12/09 07:45:12 (GMT)
SAMPLE-STOP: 1998/12/09 15:45:42 (GMT)
COUNT-START: 1998/12/09 19:17:24 (GMT)
COUNT-STOP: 1998/12/10 09:44:05 (GMT)
COUNT-TIME: 52001 .
CHECKSUM: 0A3F
COMMENTS:
Arbitrary number of lines of miscellaneous text go here. Cannot start
line with '#' sign
For example, Cell=3, Trap=B.
ARSA System Parameters OK. Next data transmission: 1998/12/11 8:00
(GMT)
#

```

The header consists of a series of “records” written one per line, with the line ending the header with a ‘#’ character. Each record begins with a field tag that defines the type of information stored in the record. Field tags are in all capital letters and are terminated with a colon (:). The field tag is the first thing on the line (excluding possible whitespace). The way that the rest of the record is interpreted depends on the particular field tag. Individual tags are described below.

In the following record descriptions:

<ws> -- stands for “one or more whitespace characters” (spaces, tabs, carriage-returns or line-feeds).

<LF> -- an ASCII line-feed character (which is the flavor of line termination we’ll use) spaces between the fields are only there for readability.

**FORMAT tag:**

<ws> FORMAT: <ws> format-name <ws> <format-version-number> <ws> <LF>

<format-name> The name of the file format (“XNG01” in the example above. The format name will not contain spaces.

<format-version-number> The version number of the format specification to which the file adheres.

**STATION tag:**

<ws> STATION: <ws> station-name <ws> <LF>

<station-name> The name of the sampling station. This field is allowed to contain spaces.

**GAMMA-CHANNELS tag:**

<ws> GAMMA-CHANNELS: <ws> num-channels <ws> <LF>

<num-channels> The total number of gamma channels in the spectrum.

**GAMMA-GAIN/OFFSET tag:**

<ws> GAMMA-GAIN/OFFSET: <ws> gain / offset <ws> <LF>

Note that the <gain> and <offset> fields are separated by a ‘/’ character and no spaces exist between the gain, the ‘/’ and the offset.

<gain> The detector gain specified in keV per channel.

<offset> The detector offset specified in keV.

**BETA-CHANNELS tag:**

<ws> BETA-CHANNELS: <ws> num-channels <ws> <LF>

<num-channels> The total number of beta channels in the spectrum. For radioxenon spectra which do not measure beta energy or do not take advantage of beta gating of the gamma spectrometer (i.e. generate only single parameter data) this field should be set to zero (‘1’).

**BETA-GAIN/OFFSET tag:**

<ws> BETA-GAIN/OFFSET: <ws> gain / offset <ws> <LF>

Note that the <gain> and <offset> fields are separated by a '/' character and no spaces exist between the gain, the '/' and the offset.

<gain> The detector gain specified in keV per channel.

<offset> The detector offset specified in keV.

#### EFFICIENCY-WINDOWS tag:

<ws> EFFICIENCY-WINDOWS: <ws> <LF>

<ws> 131m: <beta-low<sub>131m</sub>> / <beta-high<sub>131m</sub>> / <gamma-low<sub>131m</sub>> / <gamma-high<sub>131m</sub>> / <efficiency<sub>131m</sub>> <ws> <LF>

<ws> 133m: <beta-low<sub>133m</sub>> / <beta-high<sub>133m</sub>> / <gamma-low<sub>133m</sub>> / <gamma-high<sub>133m</sub>> / <efficiency<sub>133m</sub>> <ws> <LF>

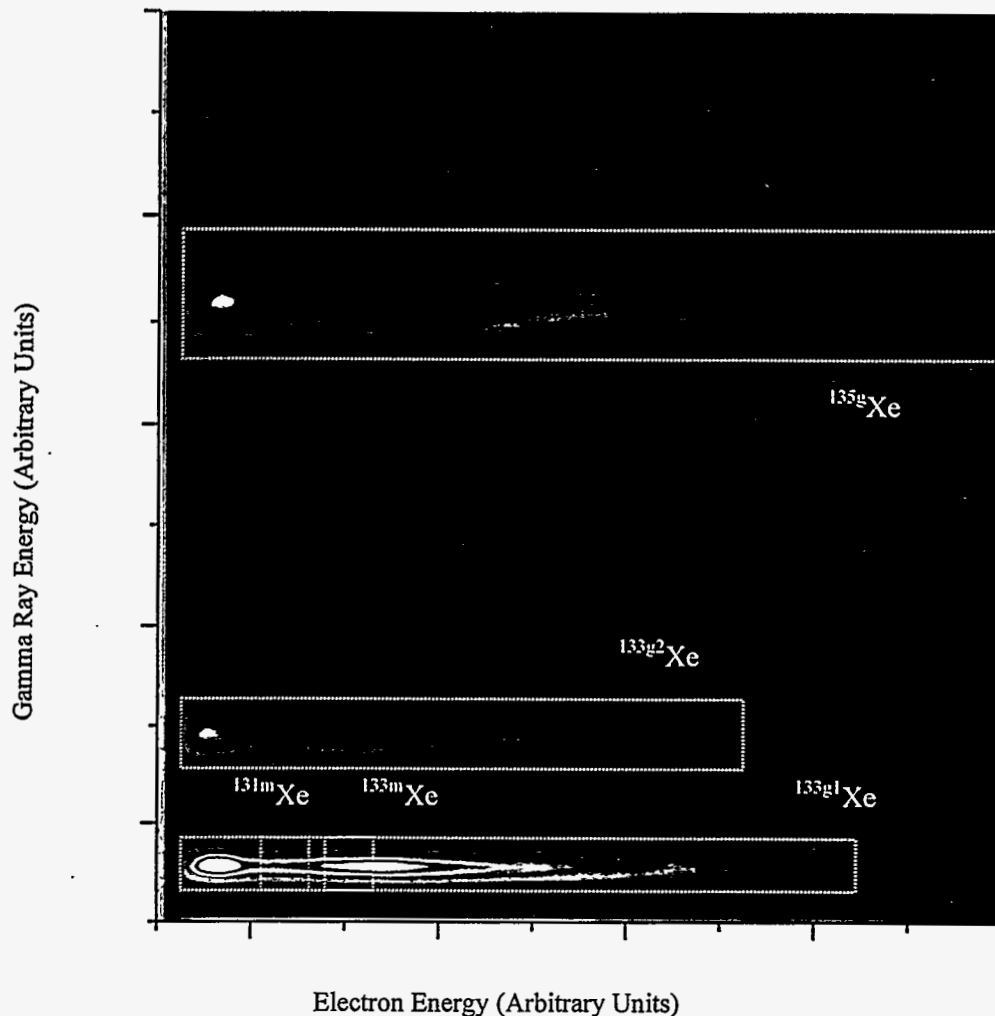
<ws> 133g1: <beta-low<sub>133g1</sub>> / <beta-high<sub>133g1</sub>> / <gamma-low<sub>133g1</sub>> / <gamma-high<sub>133g1</sub>> / <efficiency<sub>133g1</sub>> <ws> <LF>

<ws> 133g2: <beta-low<sub>133g2</sub>> / <beta-high<sub>133g2</sub>> / <gamma-low<sub>133g2</sub>> / <gamma-high<sub>133g2</sub>> / <efficiency<sub>133g2</sub>> <ws> <LF>

<ws> 135g: <beta-low<sub>135g</sub>> / <beta-high<sub>135g</sub>> / <gamma-low<sub>135g</sub>> / <gamma-high<sub>135g</sub>> / <efficiency<sub>135g</sub>> <ws> <LF>

Note that the <beta-low<sub>i</sub>>, <beta-high<sub>i</sub>>, <gamma-low<sub>i</sub>>, <gamma-high<sub>i</sub>>, and <efficiency> fields are separated by a '/' character and no spaces exist between the parameters. The <beta-low<sub>i</sub>>, <beta-high<sub>i</sub>>, <gamma-low<sub>i</sub>>, <gamma-high<sub>i</sub>> are the values in keV corresponding to corners of a two-dimensional gate or window that defines an efficiency <efficiency>. The activity of the sample is determined by dividing the number of counts over the window [after subtracting out interferences from the other radioxenons (if any) and background, and doing appropriate half-life corrections] and dividing by <efficiency> (and the count time and branching ratio corrections), depending on the isotope.

Each of the radioxenons has a single gamma-ray associated with them, except <sup>133g</sup>Xe which has two, but also have associated betas or a conversion electron associated with them. In the case of <sup>133g</sup>Xe, there are two windows whose total counts should be combined. The figure below demonstrates conceptual window placements for the radioxenons of interest.



**STABLE-XENON-VALUE** tag:

<ws> STABLE-XENON-VALUE: <ws> volume <LF>

<volume> The volume of xenon in the sample corrected to standard temperature and pressure and specified in cubic centimeters. This field will allow operators to determine the equivalent amount of air processed. The number should be in #.# format with at least one number before and after the decimal place (i.e. '1.' is not valid nor is '.996' but '1.0' and '0.996' are both valid).

**TYPE** tag:

<ws> TYPE: <ws> type-id <ws> <LF>

<type-id> One of the strings 'SAMPLE', 'CALIBR', or 'BACKGR' (without the quotes). SAMPLE means the file contains a spectrum of an actual xenon sample, CALIBR means the spectrum of a calibration source, and BACKGR denotes a background spectrum taken without a sample.

**DATE** tag:

<ws> DATE: <ws> date-string <ws> <LF>



**<date-string>** A string in the following format (excluding the quotes) “YYYY/MM/DD HH:MM:SS (GMT)” which gives the date and time the file was generated. YYYY, MM, and DD represent the year month and day respectively and HH, MM, SS are hour, minute, and second values. Note that the year is a 4-digit number (i.e., 1998) and all other numbers are 2-digits long and padded with a zero (‘0’) on the left if less than 10 (i.e. 07 or 14). The ‘month’ and ‘day’ both start counting at ‘01’ (i.e., January 1st is MM=01, DD=01) and ‘hour’, ‘minute’ and ‘second’ begin counting at zero. HH assumes a 24-hour clock and has a valid range of 00 to 23 with midnight = 00. Time and date are reported as Greenwich Mean Time (GMT).

**SAMPLE-START tag:**

**<ws> SAMPLE-START: <ws> date-string <ws> <LF>**

**<date-string>** A string in the following format (excluding the quotes) “YYYY/MM/DD HH:MM:SS (GMT)” which gives the date and time the system started accumulating the sample (not the count data). YYYY, MM, and DD represent the year month and day respectively and HH, MM, SS are hour, minute, and second values. Note that the year is a 4-digit number (i.e. 1998) and all other numbers are 2-digits long and padded with a zero (‘0’) on the left if less than 10 (i.e. 07 or 14). The ‘month’ and ‘day’ both start counting at ‘01’ (i.e. January 1st is MM=01, DD=01) and ‘hour’, ‘minute’ and ‘second’ begin counting at zero. HH assumes a 24 hour clock and has a valid range of 00 to 23 with midnight = 00. Time and date are reported as Greenwich Mean Time (GMT).

**SAMPLE-STOP tag:**

**<ws> SAMPLE-STOP: <ws> date-string <ws> <LF>**

**<date-string>** A string in the following format (excluding the quotes) “YYYY/MM/DD HH:MM:SS (GMT)” which gives the date and time the system stopped accumulating the sample (not the count data). YYYY, MM, and DD represent the year month and day respectively and HH, MM, SS are hour, minute, and second values. Note that the year is a 4-digit number (i.e. 1998) and all other numbers are 2-digits long and padded with a zero (‘0’) on the left if less than 10 (i.e. 07 or 14). The ‘month’ and ‘day’ both start counting at ‘01’ (i.e. January 1st is MM=01, DD=01) and ‘hour’, ‘minute’ and ‘second’ begin counting at zero. HH assumes a 24 hour clock and has a valid range of 00 to 23 with midnight = 00. Time and date are reported as Greenwich Mean Time (GMT).

**COUNT-START tag:**

**<ws> COUNT-START: <ws> date-string <ws> <LF>**

**<date-string>** A string in the following format (excluding the quotes) “YYYY/MM/DD HH:MM:SS (GMT)” which gives the date and time the system started counting the sample. YYYY, MM, and DD represent the year month and day respectively and HH, MM, SS are hour, minute, and second values. Note that the year is a 4-digit number (i.e., 1998) and all other numbers are 2-digits long and padded with a zero (‘0’) on the left if less than 10 (i.e. 07 or 14). The ‘month’ and ‘day’ both start counting at ‘01’ (i.e. January 1st is MM=01, DD=01) and ‘hour’, ‘minute’ and ‘second’ begin counting at zero. HH assumes a 24 hour clock and has a valid range of 00 to 23 with midnight = 00. Time and date are reported as Greenwich Mean Time (GMT).

**COUNT-STOP tag:**

**<ws> COUNT-STOP: <ws> date-string <ws> <LF>**

**<date-string>** A string in the following format (excluding the quotes) “YYYY/MM/DD HH:MM:SS (GMT)” which gives the date and time the system stopped counting the sample. YYYY, MM, and DD represent the year month and day respectively and HH, MM, SS are hour, minute, and second values. Note that the year is a 4-digit number (i.e., 1998) and all other numbers are 2-digits long and padded with a zero (‘0’) on the left if less than 10 (i.e. 07 or 14). The ‘month’ and ‘day’ both start counting at ‘01’ (i.e. January 1st is MM=01, DD=01) and ‘hour’, ‘minute’ and ‘second’ begin counting at zero. HH assumes a 24 hour clock and has a valid range of 00 to 23 with midnight = 00. Time and date are reported as Greenwich Mean Time (GMT).

**COUNT-TIME tag:**

<ws> COUNT-TIME: <ws> seconds <LF>

<seconds> The total duration of the sample count in seconds.

**CHECKSUM tag:**

<ws> CHECKSUM: <ws> checksum-value <ws> <LF>

<checksum-value> A four character ASCII encoded hexadecimal number that is generated by summing the ASCII values of all non-whitespace characters in the file into a 2-byte integer. This field is used to detect file corruption.

**COMMENTS tag:**

<ws> COMMENTS: <ws> <LF>

<ws> Comment 1 <LF>

<ws> Comment 2 <LF>

<ws> Comment 3 <LF>

.

.

.

<ws> Comment n <LF>

<ws> # <LF>

The comments section is allowed to contain an arbitrary number of lines of comments. The end of the data header is signified by the '#' character.

**FORMAT DETAILS: DATA SECTION**

The histogram data can be thought of as a 2-dimensional array of integers which we'll represent as

hist[Ng,Nb]

where (Ng) is the number of gamma channels and (Nb) is the number of beta channels (both of these are specified in the header).

hist[0,0] = lowest gamma energy, lowest beta energy

hist[Ng-1,Nb-1] = highest gamma energy, highest beta energy

The array is stored row by row as a series of integers separated by a comma and whitespace. The numbers are ordered as follows

hist[0,0], hist[1,0] hist[2,0] ... hist[Ng-2,0] hist[Ng-1,0]

hist[0,1] hist[1,1] hist[2,1] ... hist[Ng-2,1] hist[Ng-1,1]

hist[0,2] hist[1,2] hist[2,2] ... hist[Ng-2,2] hist[Ng-1,2]

.

.

.

```
hist[0,Nb-2] hist[1,Nb-2] hist[2,Nb-2] ... hist[Ng-2,Nb-2] hist[Ng-1,Nb-2]  
hist[0,Nb-1] hist[1,Nb-1] hist[2,Nb-1] ... hist[Ng-2,Nb-1] hist[Ng-1,Nb-1]
```

Note that the only requirement is that the individual numbers be separated by one or more whitespace characters (tabs, spaces, carriage-returns, or linefeeds). This leaves the implementer free to organize the data however they'd like. For instance, the minimal file size can be achieved by separating the numbers by single spaces only and including no line-feeds. The resulting file would be nearly impossible to interpret with a text editor though. To generate a file that was more friendly to human readers the implementer could, for example, choose to line the numbers up on 5-character columns, add a line-feed after every 16th number and add two line-feeds at the end of each "hist[0,n] .. hist[Ng-1,n]" row. Keep in mind however that the additional formatting whitespace would increase the size of the file by  $2 \frac{1}{2}$  times.

## Conclusions

A data format has been presented appropriate for 2 parameter data. This format is consistent with the data typical for Automated Radioxenon Sampler-Analyzer (ARSA) data, as well as the type of data expected soon from the Swedish xenon systems<sup>1</sup>. This format should be considered a "straw man" to spur discussions on how this data fits into the current IMS data structure.

## References

1. Anders Ringbom, private communication.

## DISTRIBUTION

No. of Copies

## OFFSITE

1	Leslie Casey US DOE/HQ 1000 Independence Avenue, SW Washington, DC 20585	1	Pres Herrington Sandia National Laboratories Dept. 5736/MS 0655 PO Box 5800 Albuquerque, NM 87185-0655
1	Marshall Billingslea Senate Foreign Relations Committee 450 Dirksen Senate Office Building Washington, DC 20510	1	Pujing Pan Radiation Protection Bureau, Health Canada 775 Brookfield Road Ottawa/Ontario/K1A 1C1
1	Richard Morrow Arms Control and Disarmament Agency/IVI 320 21 <sup>st</sup> St. NW Room 5741 Washington, DC 20451	1	Anders Ringbom Defense Research Establishment S-172 90 Stockholm, Sweden
1	Dr. Joachim Schulze PTS/PrepCom/CTBTO Vienna International Centre, Room E0764 PO Box 1200 A-1400 Vienna, Austria		
1	Jim Dishaw CIA 4T31 NHB Washington, DC 20501		
1	Ralph Keyser Sandia National Laboratories Dept. 6532/MS 1138 PO Box 5800 Albuquerque, NM 87185-12238		

## ONSITE

23	PNNL  Ray Warner TW Bowyer TR Heimbigner JI McIntyre AD McKinnon PL Reeder ER Wittinger Donna Schneider (10) Information Release (7)
----	--