Results of the Automated and Interactive Radionuclide Reports comparison between International Data Centre (IDC) and baseline reports from National Data Centres (NDCs)

D. Stasic, S. Laban

Comprehensive Nuclear-Test-Ban Treaty Organization (CTBTO)

CTBTO PREPARATORY COMMISSION

INTRODUCTION AND MAIN RESULTS

The National Data Centres (NDCs) reports are, for the first time, used to produce baseline reports for comparison with the IDC products for three quality parameters: (1) number of peaks detected (2) number of correctly identified relevant radionuclides, and (3) activity concentration of the quantified radionuclides. The metrics used for nuclide quantification accuracy comparison are absolute percent difference between the IDC and the NDCs activity concentration and zeta score (ζ) in line with ISO 13528:2022(E).

This poster will demonstrate the methodology used for the quality assessment of the IDC radionuclide products and summarize the main findings for particulate samples.

Results of the Automated and Interactive Radionuclide Reports comparison between International Data Centre (IDC) and baseline reports from National Data Centres (NDCs)

D. Stasic, S. Laban

P3.2-865

Introduction

The 2024 Experiment took place during the period 16–27 September 2024. This experiment was the second in the new cycle as part of the ongoing testing and evaluation activities under the PTS performance monitoring and testing framework Rev.1 (CTBT/PTS/INF.1046/Rev.1). During 2024 Experiment, Validation tests (VTs) related to the automatic processing and interactive analysis of IDC RN products were retested. The quality of IDC produced Automatic Radionuclide Reports (ARRs) and Reviewed Radionuclide Reports (RRRs) in 2024 Experiment was evaluated by comparing IDC products with baseline reports generated by participating NDCs against preset criteria derived by IDC

Baseline reports

National Data Centres (NDCs) participated in 2024 Experiment by performing Automated and Interactive analysis of Radionuclide Spectral Data and produced both the Automated Radionuclide Reports (ARR) and the Reviewed Radionuclide Reports (RRR) for a selection of particulate and noble gas sample spectra. The NDC reports were used to produce the baseline reports for comparison with the IDC products for three quality parameters: (1)* number of peaks detected (2) number of correctly identified relevant radionuclides, and (3) activity concentration of the quantified radionuclides. The metrics used for nuclide quantification accuracy comparison are percent difference (%D) between the IDC and the baseline activity concentration and zeta (ζ) score as defined in ISO 13528:2022(E). This poster will demonstrate the methodology used for the quality assessment of the IDC radionuclide products and summarize the main findings for particulate samples.

*peak detection and identification apply only to particulate systems

Quality metrics

Item	Acceptance Criteria	Metric	Target
(1)	The number of peaks with Peak Significance ≥1 identified in the IDC ARR and baseline ARR shall not differ by more than 20%. (10% for RRRs)	The percentage of particulate ARRs with % Difference in the number of peak detections with Peak Significance(1) ≥1 between IDC ARRs the baseline ARRs ≤ 20%. (≤ 10% for RRRs)	95%(2) for ARRs 98% for RRRs
(2)	The true positive rate(4) of identification of relevant radionuclides shall be at least **95% for ARRs** (with not more than 5% false positives) **and 98% for RRRs** (with not more than 2% false positives).	The percent true positive rate of relevant nuclide identifications and percent false positive rate of relevant nuclide identifications for the ARR (RRR) particulate population of samples.	95% TP, 5% FP for ARR 98% TP, 2% FP for RRR
(3)	The activity concentration values of radionuclides quantified in IDC products shall meet the following criteria, using the baseline values as reference: • ARRs: |%D| ≤ 10%, Zeta ≤ 2 • RRRs: |%D| ≤ 5%, Zeta ≤ 2	The percentage of particulate ARRs with the concentration value that meets the following criteria, considering the baseline values as a reference: • ARRs: |%D| ≤ 10%, Zeta ≤ 2 • RRRs: |%D| ≤ 5%, Zeta ≤ 2	95% for ARRs 98% for RRRs

Scope of Work

PARSED	IDC sent	NDC1	NDC2	NDC3	NDC4	NDC5	NDC6	NDC7
ARR Particulate	214	30	210	208	0	206	195	200
RRR Particulate	214	11	214	213	0	203	38	0

IDC Risk level is 0.001% (k_α=4.2649) where most NDCs use different risk level. This is also the case with different algorithms for calibrations, peak detection and nuclide identification between IDC and NDC and different nuclear libraries. Differences in the processing parameters between IDC and NDCs introduce discrepancies in the results.

The raw data analysed by NDCs has been divided into two groups:

1. 100 SPHDs were chosen from the samples with detected anthropogenic nuclides: 30 Particulate, 35 SPALAX and 35 Beta/Gamma noble gas.
2. 3 days of data acquired during 2024 Experiment: 18th, 21st and 22nd September 2024.

Special DB schema to hold NDC results, currently contains **40 Tables**:

- 1 IDC Samples Definition Table
- 1 IDC Stations definition Table
- 22 ARR/RRR Tables for raw Particulates and Spalax
- 12 ARR/RRR BG Tables for raw Sauna reports
- 4 revised/unified results tables for NG reports

CTBTO PREPARATORY COMMISSION

Results of the Automated and Interactive Radionuclide Reports comparison between International Data Centre (IDC) and baseline reports from National Data Centres (NDCs)

D. Stasic, S. Laban

P3.2-865

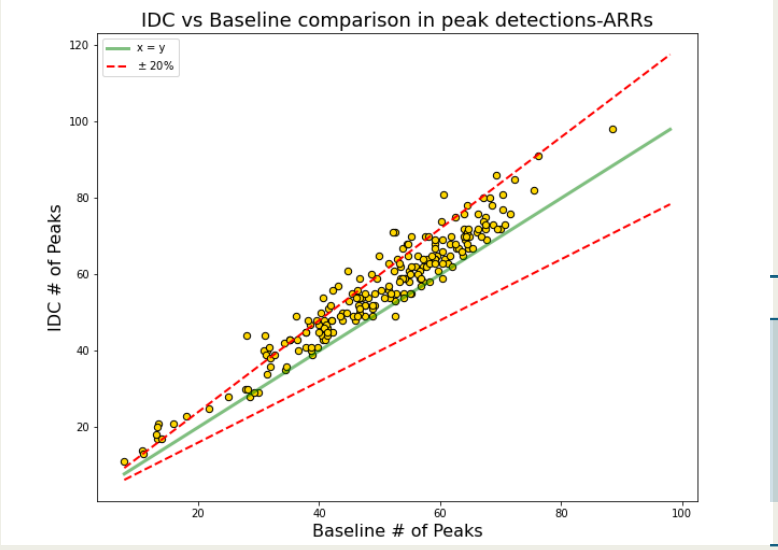

Figure 1. IDC vs Baseline comparison in peak detections-ARRs (particulate) 77% of the samples meet the criteria

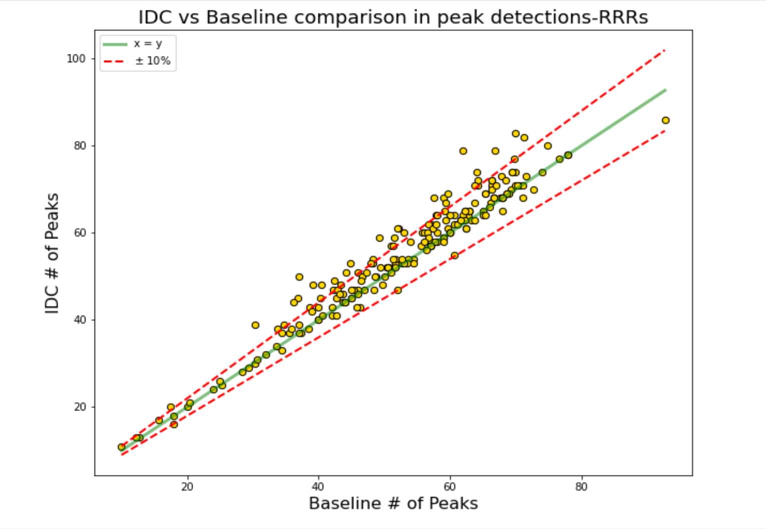

Figure 2. IDC vs Baseline comparison in peak detections-RRRs (particulate) 80% of the samples meet the criteria

Metric 1: Number of peak detections

The baseline number of peak detections with a Peak Significance ≥1 is calculated as a robust average of the number of peaks with Sig ≥ 1 reported by NDCs for the same sample. Peak Significance = Peak Area/ Lc (kα). The results are shown in Figure 1 and 2.

Metric 1	Result	Target
The percentage of particulate samples with % Difference ≤ 20% in the number of peak detections with Peak Significance(1) ≥1 between baseline ARR and the IDC ARR.	77%	95%
RRR: %Difference ≤ 10%	80%	98%

Metric 2: Percent of correct identifications

	NR	TP	FN	FP	TN	TPR	FPR
ARR	128	13	21	98	17416	38.2%	0.6%
RRR	162	13	0	62	17473	100.0%	0.4%

Metric 2	Result	Target
The percent true positive rate of relevant nuclide identification and percent false positive rate of relevant nuclide identification for the ARR particulate population of samples.	38.2%TP 0.6% FP	95%TP 5%
Same metric as above for RRR	100% 0.4%	98%TP 2% FP

- $TPR = \frac{TP}{TP+FN}$ $FPR = \frac{FP}{FP+TN}$
- $TPR = \frac{TP*100}{TP+FN}$, TP+FN=13+21
- Max number $_{\text{id-ed relevant nuclides}}$ = TP+FN+FP+TN
- Max number = 82x214
- TN = (Max number(82x214))– (TP+FN+FP)
- $FPR = \frac{FP*100}{17548-TP-FN}$

NR- no relevant nuclides found

The baseline report contains only relevant nuclides identified in at least 50% of the NDC generated reports for the same sample

SID	NoRelevant	CE-144	CS-137	GA-72	RU-103	I-131	I-133	PB-203	ZR-95	NA-24	BA-140	SB-120	NB-95	RB-84	SB-126
7139206		FP													
7198806			TP	FP											
7240519		FP			FP										
7589441			TP			TP	FN								
7647390	NR														
7650953								FP	FP						
7657763										TP	FP	FP			
7660823										FP			FP	FP	
7665868			TP			TP									
7668548			FN												
7670354								FP							FP
7675332			TP										FP		
7692316										FP			FP	FP	
7698029	NR														
7720222			FN										FP		
7731668			FP											FP	
7748616	NR														
7760662										FP					
7770801	NR														
7777686			TP												
7811412								FP		FP	FP			FP	FN
7814989	NR														
7818161															
7827438															
7870539										TP	FP				
7876072	NR														

214

82

Figure 3. CTBT Relevant Nuclide truth table (only a segment is shown)

CTBTO PREPARATORY COMMISSION

Results of the Automated and Interactive Radionuclide Reports comparison between International Data Centre (IDC) and baseline reports from National Data Centres (NDCs)

D. Stasic, S. Laban

P3.2-865

Metric 3: Nuclide quantification accuracy

The standard metrics are used to calculate nuclide quantification accuracy, also used in the network QA/QC program (refer to IDC-QMS-PLN-850, Particulate and Noble Gas Quality Assurance Plan) and originate from ISO 13528:2022(E) (latest available version) Statistical Methods for Use in Proficiency Testing by Interlaboratory Comparison.
The test for accuracy of nuclide quantification is using the absolute percent difference (%D) between IDC and baseline activity concentration and zeta (ζ) score defined as:

$$\%D = \left|\frac{A_{NDC} - A_{IDC}}{A_{NDC}}\right| \times 100$$

$$\text{Zeta score } (\zeta) = \frac{|Ac_{NDC} - Ac_{IDC}|}{\sqrt{u^2{}_{NDC} + u^2{}_{IDC}}}$$

Note!: The NDCs were not subtracting the BLANK filter measurements from the SAMPLE measurements and therefore, the IDC concentrations had to be recalculated to obtain the values before the blank subtraction.
Those were not available directly from IDC database but had to be calculated separately as well as the concentration uncertainties. The activity concentrations of quantified nuclides were calculated before the BLANK measurement subtraction and compared to the baseline concentrations calculated as a robust average of participating NDCs' concentration. The results for ARRs are shown in Figure 4 and 5 and for RRRs in Figures 6 and 7.

Metric 3	Result	Target
The percentage of particulate **ARRs** with the activity oncentration value that meets criteria of **|%D| ≤ 10%, Zeta-Score ≤ 2**, considering the baseline values as a reference ARRs:	|%D| Be-7: 97.2% Zeta-Score BE-7: 99.5% |%D|CTBT Rel:77 % Zeta-Score CTBT Rel: 100%	95%
The percentage of particulate **RRRs** with the concentration value that meets the criteria of **|%D| ≤ 5% and Zeta-Score ≤ 2**, considering the baseline value as a reference.	|%D| Be-7: 91.6% Zeta-Score BE-7: 99% |%D|CTBT Rel: 62% Zeta-Score CTBT Rel:100%	98%

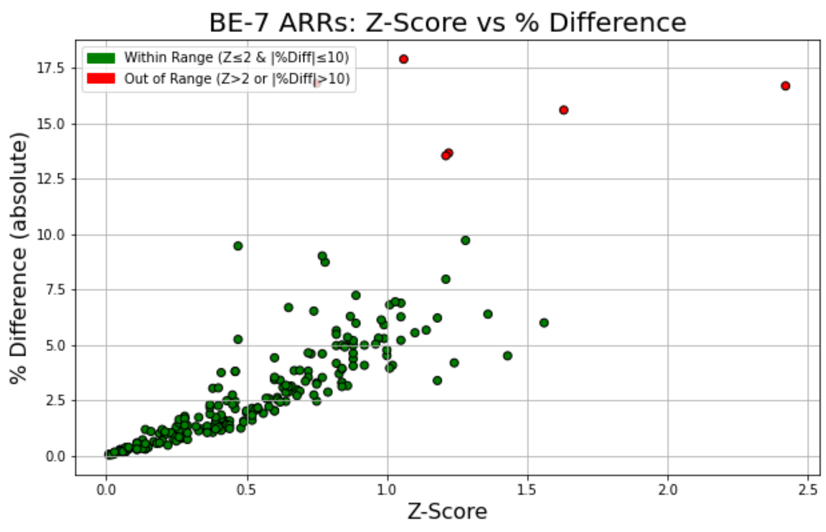

Figure 4. % Difference(abs) vs Z-Score in ARR activity concentrations of Be-7

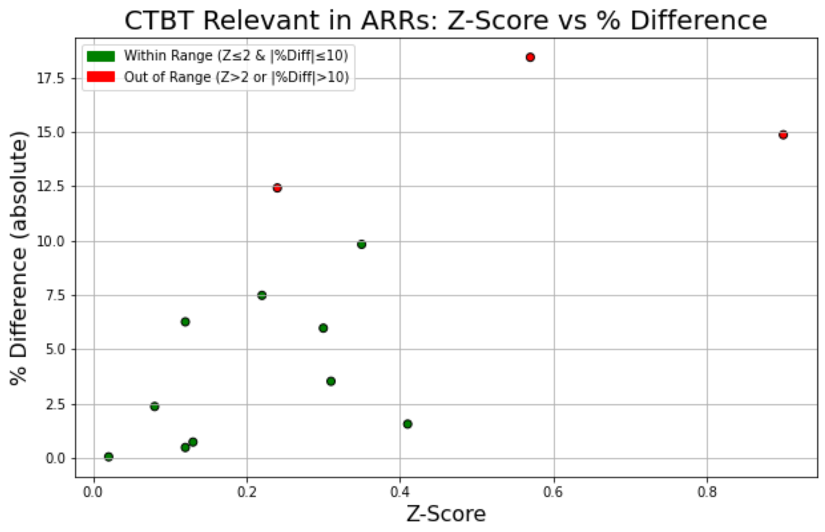

Figure 5 % Difference(abs) vs Z-Score in ARR activity concentrations of CTBT relevant radionuclides

CTBTO PREPARATORY COMMISSION

SnT 2025
8 SEPTEMBER ONLINE DAY
9 TO 12 SEPTEMBER AT HOFBURG PALACE, VIENNA & ONLINE
CTBT: SCIENCE AND TECHNOLOGY CONFERENCE

Results of the Automated and Interactive Radionuclide Reports comparison between International Data Centre (IDC) and baseline reports from National Data Centres (NDCs)

D. Stasic, S. Laban

P3.2-865

Metric 3-RRRs

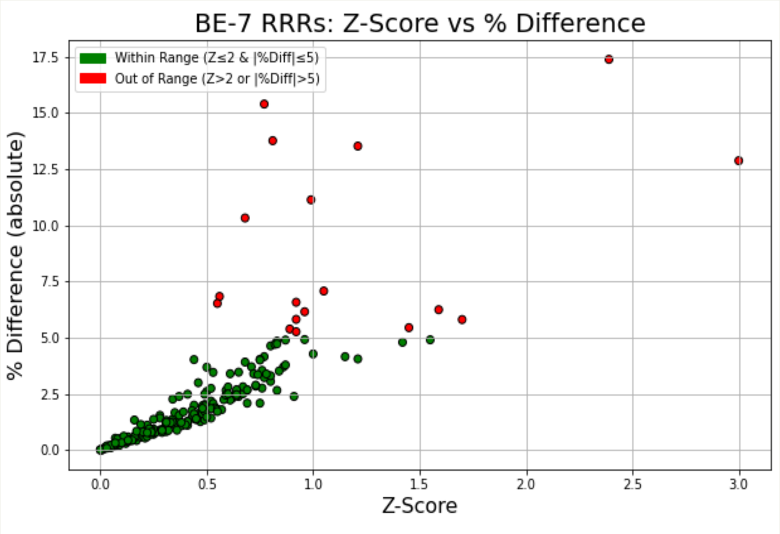

Figure 6. % Difference(abs) vs Z-Score in RRR activity concentrations of Be-7

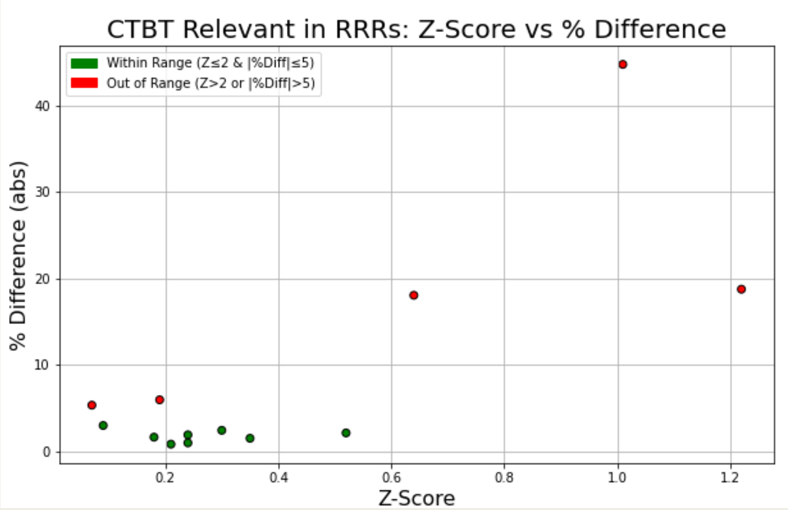

Figure 7 % Difference(abs) vs Z-Score in RRR activity concentrations of CTBT relevant radionuclides

Lessons Learned

CHALLENGES

Complex Parsing, new Database tables and Data formatting & Management

- Extracting and storing data from different NDC reports required significant effort
- Difficulty with missing sample reports, processing of bad samples, reports with missing information and reports containing errors

Diverse Report Formats:

- Each of the **seven NDC provided reports** with **unique format (i.e. txt, html, xlsx/csv)**, since they are generated by NDC specific software, requiring custom parsing scripts
- Sometimes different format for ARR and RRR from the same NDC

BENEFITS

- Regular feedback and communication with NDCs regarding missing information and fixing errors in the reports may help improve their software
- The developed framework for comparison of IDC RN products with NDCs' baseline will be used for future comparisons

Conclusions

- The IDC software successfully identified all relevant radionuclides in the provided data set, with a false negative rate of 0%, i.e. no relevant radionuclides were missed
- The IDC reports a slightly higher number of peaks compared to the NDCs' baseline and has a higher false positive rate
- Differences in the processing parameters between IDC and NDCs introduce discrepancies in the results
- The legacy targets need to be revised and adapted to the new tests
- The lack of ground truth data introduces significant uncertainties in the comparison between IDC data and the baseline
- The ideal scenario would be to have a set of synthetic spectra (as the ground truth) analysed by participating NDC using different software and compare the results with the IDC

References

1. CTBT/PTS/INF.1741 (July 2025). Technical Report on the 2024 Experiment. Provisional Technical Secretariat of the CTBTO Preparatory Commission, Vienna, Austria.
2. IDC-QMS-PLN-001.Rev.5.4 (July 2024). Validation and Acceptance Test Plan. Provisional Technical Secretariat of the CTBTO Preparatory Commission, Vienna, Austria.
3. IDC-QMS-PLN-850 (July 2018). Particulate and Noble Gas Quality Assurance Plan. Provisional Technical Secretariat of the CTBTO Preparatory Commission, Vienna, Austria.
4. ISO 13528:2022(E) (2022). Statistical Methods for Use in Proficiency Testing by Interlaboratory Comparison. International Organization for Standardization.
5. ECS/PRES/WGB-62/RNEG/2 (February 2023). RN Software Benchmarking - Part 2 (Presentation to Radionuclide Expert Group during Working Group B 62). Provisional Technical Secretariat of the CTBTO Preparatory Commission, Vienna, Austria.

CTBTO PREPARATORY COMMISSION