

Evaluating Cs-137 Analytical Variations in IMS Network Particulate Samples: Insights from Ten Years of IDC and Radionuclide Laboratories Results

Rodrigo Villarreal, Martina Rozmaric, Marina Nizamska, Felix Pino, Seokryung Yoon, Jana Meresova, Eric Nguelem, Jonathan Bare^[1]

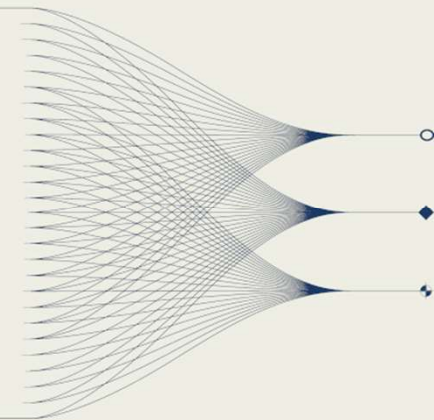
CTBTO IDC Division, PO Box 1200, 1400 Vienna, Austria

[1] Former Lead Analyst IDC/MDA



..... INTRODUCTION AND MAIN RESULTS

The International Monitoring System (IMS) is a unique global network designed to monitor compliance with the Comprehensive Nuclear-Test-Ban Treaty (CTBT). Complementary to three waveform technologies, the 80-station radionuclide network provides continuous worldwide monitoring. This radionuclide network is further supported by 16 specialized laboratories, which offer independent expert analysis of IMS samples upon request. This work presents a comparative assessment of Cs-137 analysis results performed by the International Data Centre (IDC) and Provisional Technical Secretariat-certified radionuclide laboratories over the past decade. Cs-137 was selected for its relevance to monitoring, its long half-life (thus mitigating decay-related uncertainties), and because initial difference tests obtained from the data evaluation revealed notable discrepancies. Using both difference tests and zeta scores, potential causes of these inconsistencies, including peak interference, baseline underestimation, and background subtraction methods were investigated. The obtained results underscore the need for ongoing analytical refinements to ensure reliable radionuclide measurements in support of effective verification activities.



.....
DISCLAIMER: The views and opinions expressed in this presentation are those of the authors and do not necessarily represent official policy or position the Comprehensive Nuclear-Test-Ban Treaty Organization (CTBTO).

Rodrigo Villarreal, Martina Rozmaric, Marina Nizamska, Felix Pino, Seokryung Yoon, Jana Meresova, Eric Nguelem, Jonathan Bare

Introduction and Motivation

Methods

P4.1-874

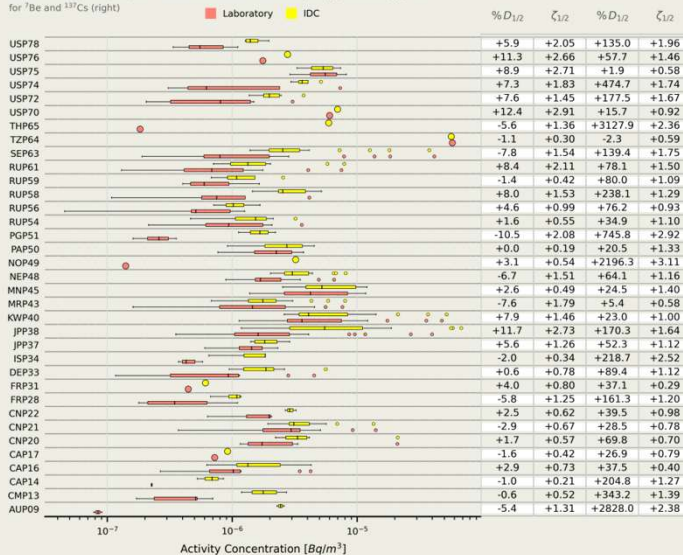
Why Cs-137?

Cesium-137 is a long-lived fission product (half-life ~30 years) of high relevance for treaty verification. Its presence in a sample can always be confirmed by laboratory analysis in non-split samples.

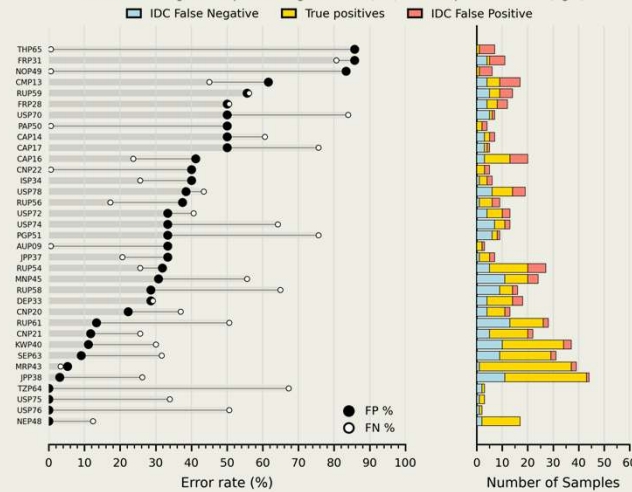
Study Dataset

We analysed 277 Category A samples (true positives where both IDC and the Laboratory confirmed Cs-137).

Station-Level Distributions of ¹³⁷Cs Activity Concentrations: IDC vs Laboratory
Boxplots on log scale (left) with station-level median difference (%D_{1/2}) and zeta (ζ_{1/2}) metrics for ⁷Be and ¹³⁷Cs (right)



¹³⁷Cs Detection Error Rates: IDC vs. IMS Laboratory (2015-2025)
Per-station ranking of false positive/negative rates (left) with sample size context (right)



Laboratory as Ground Truth

Laboratories employ long counting times (up to 7 days) and advanced equipment such as ultra-low-background detectors, shielding, and cosmic veto systems. These conditions ensure high-confidence measurements.

Key Findings

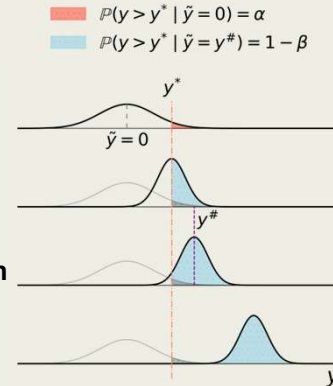
Comparison of IDC and laboratory results shows systematic differences, highlighting opportunities to refine analysis methods. IDC results show a strong positive bias (**median difference = +63%**), with extremes exceeding 1000%. Some stations exhibit very high false-positive rates (**up to 86%**).

Reanalysis approach:

Adapted elements of the manual analyst workflow; peak areas fitted using weighted counts (channel uncertainties as weights).

Decision limits:

Recalculated **decision threshold (y*)** and **detection limit (y#)** following a partial ISO 11929^[1] implementation (relative calibration factor **w** uncertainty < 8%).



$$y^* = k w \sqrt{n_0 + n_b + u^2(n_0) + u^2(n_b)}$$

$$y^\# \simeq k^2 w + 2y^*$$

Additional variance terms: Included baseline (**n₀**) variance and time-scaled background contribution (**n_b**) with its variance, both ignored in the current IDC algorithm.

Baseline statistics: Investigated the quality of the baseline fit using the **reduced χ² test** in the **661.7 ± 20 keV window**, ensuring proper modelling of statistical fluctuations around the Cs-137 peak region..

Rodrigo Villarreal, Martina Rozmaric, Marina Nizamska, Felix Pino, Seokryung Yoon, Jana Meresova, Eric Nguelem, Jonathan Bare

P4.1-874

Baseline Statistics and Possible Interferences

Baseline modelling

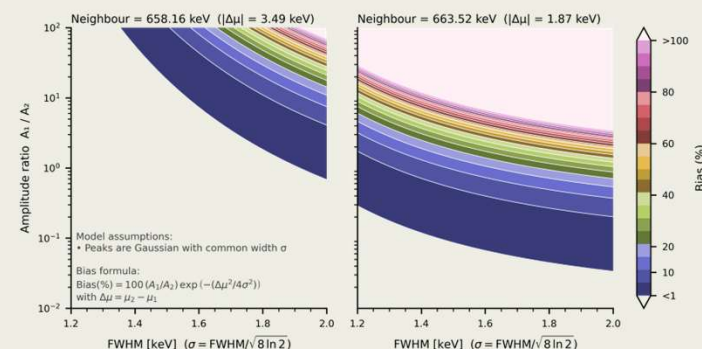
To improve reliability in decision limits, baseline statistics must be included, both their variance and the reduced χ^2 goodness-of-fit. These provide a measure of how well the baseline represents the continuum and ensure that baseline fluctuations contribute properly to decision limits. In addition, a visualization tool to evaluate the standardized residuals (white triangles) as shown in the first plot will assist the analyst in characterizing the baseline more accurately.

In this study, baseline quality was assessed through **standardized residuals** in non-peak regions (white triangles in the plot), within the **661.7 ± 20 keV window**. Overall, the baseline fits the data well, with a **median reduced χ^2 of 1.11**. The **variance in counts is ~12%**, and for this dataset it can contribute up to **30% to the decision limits**. This effect grows if peaks in the region are missed.

Interference study

Potential interferences near the Cs-137 line were also examined. The closest feature, the **Compton edge of TI-208 860.5 keV peak at 663.5 keV**, was tested under different resolutions and area ratios; it was not observed in the data, and any possible bias is **<10%**. The **TI-208 x-ray/gamma sum peak (583.19 ± 74.97 keV) at 658.2 keV** is farther away and does not introduce significant bias, even if unmodeled.

Bias on ^{137}Cs main peak area from unmodeled neighbouring line (contour comparison)
Filled contours of expected bias (%) over detector FWHM and amplitude ratio
 $A_2 = ^{137}\text{Cs}$ main peak (661.65 keV; μ_2); A_1 = neighbouring line (658.16 or 663.52 keV; μ_1)



Background subtraction

Whenever a filter batch is changed at a station, a **blank measurement** is performed, normally over **72h** (minimum^[2] 24h for paired blanks). Detector background is measured only during **station certification**, but additional checks may be needed in extraordinary cases of contamination to isolate background contributions and guide mitigation actions. Blank subtraction is applied **line by line**, scaled to the acquisition duration of the sample measurement. Under normal circumstances, **Cs-137 is not expected** in blank filters. However, the **scaled peaked background and its variance are not yet implemented** in the decision-limit algorithms, an aspect addressed in the next section.

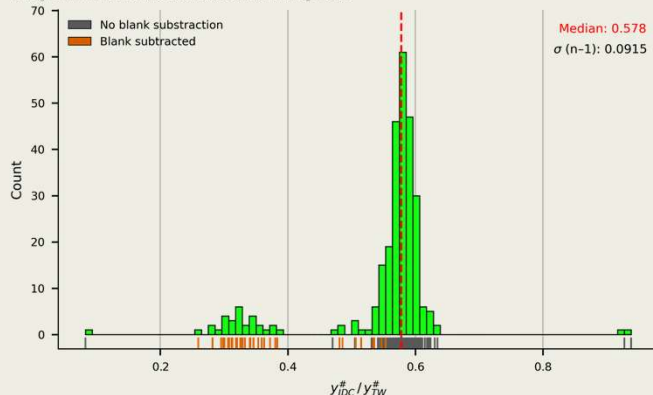
Rodrigo Villarreal, Martina Rozmaric, Marina Nizamska, Felix Pino, Seokryung Yoon, Jana Meresova, Eric Nguemem, Jonathan Bare

P4.1-874

Comparison and assessment of detection limits

The current IDC algorithm for determining detection limits does not account for **baseline variance** or **background subtraction terms**.

Distribution of MDC Ratio (IDC vs This Work)
Histogram with median (red dashed), $\sigma(n-1)$, and rug marks



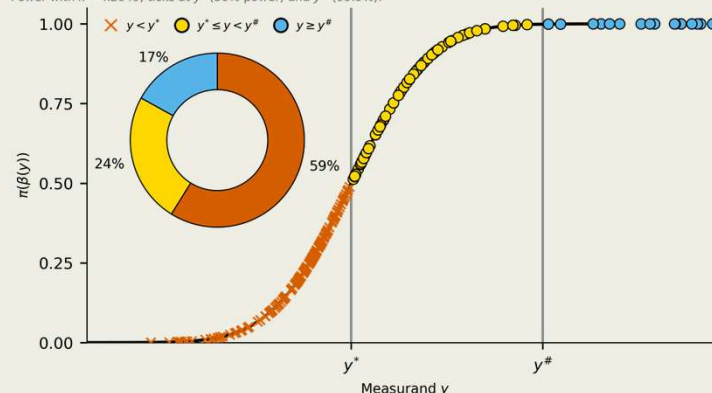
In this work, these contributions were included, and the comparison shows detection limits (MDCs) about **a factor of two higher** than IDC values, even without background subtraction.

Further review revealed an **inconsistency in ROI definitions**: IDC applies an optimized ROI of $1.25 FWHM_C$ when calculating decision thresholds (per De Geer^[3]), but peak areas are integrated over $2.5 FWHM_C$.

This mismatch results in **underestimated detection limits**, creating a false impression of sensitivity and contributing to elevated false positive rates. Importantly, this issue is not limited to **Cs-137** but applies to **all nuclides**. When harmonized ROIs are used for both detection limits and peak areas, results show that at the current risk level **~60% of Cs-137 peak areas fall below the decision threshold**. In other words, areas without statistical significance are being reported as detections.

Lessons learned: A consistent treatment of ROIs, along with the inclusion of baseline and background contributions, is essential for realistic detection limits and reliable reporting. Addressing these aspects in future IDC software will strengthen confidence in reported results and reduce false positives.

Gaussian Power Curve with Data Overlay ($\alpha = 0.001$)
Power with $k = 4.2649$; ticks at y^* (50% power) and $y^\#$ (99.9%).



Conclusions

The review of **Cs-137 comparisons** revealed key opportunities to improve baseline treatment, detection limit estimation, and reporting workflows. Incorporating baseline variance, harmonizing ROI definitions, and aligning algorithms with established standards will make results more robust and transparent.

These measures will:

- Strengthen data integrity and compliance with rules.
- Reduce false positives and improve confidence in reported detections.
- Provide analysts with clearer tools for accurate and timely reporting.

Some open points, such as inhomogeneous activity, require further study. Addressing these opportunities together will **enhance the reliability of results and reinforce confidence in the data we deliver**.

References

1. ISO (2019). ISO 11929-1:2019 – Determination of the characteristic limits (decision threshold, detection limit and limits of the coverage interval) for measurements of ionizing radiation – Fundamentals and application – Part 1: Elementary applications. International Organization for Standardization.
2. CTBT/WGB/TL-11, 17/18/Rev.7 Operational Manual for Radionuclide Monitoring and the International Exchange of Radionuclide Data
3. De Geer, L.-E. (2004). Currie detection limits in gamma-ray spectroscopy. Applied Radiation and Isotopes, 61(2–3), 151–160. <https://doi.org/10.1016/j.apradiso.2004.03.037>