GR 25:04

Quality Assessment of Updated Gamma Spectroscopy Methods

Measurement and Analysis Procedures in the

Icelandic Radiation Safety Authority Monitoring Program

GEISLAVARNIR RÍKISINS
ICELANDIC RADIATION SAFETY AUTHORITY

Quality Assessment of Updated Gamma Spectroscopy Methods

Measurement and Analysis Procedures in the Icelandic Radiation Safety Authority Monitoring Program

Authors: Henrik Öberg, Gísli Jónsson, Jónína Guðjónsdóttir, Pernille Ahlmann Jensen, Kjartan Guðnason.

Geislavarnir ríkisins
Rauðarárstíg 10
105 Reykjavík
sími: 440 8200
gr@gr.is
www.gr.is

Útgáfa 25:04

ISBN 978-9935-9861-0-8

Reykjavík, júní 2026

Table of Contents

Summary

Between 2020 and 2021, the Icelandic Radiation Safety Authority (IRSA) monitoring programme implemented several methodological updates. Measurement containers were replaced: 200 ml containers used for media such as milk powder were upgraded to 350 ml, and Marinelli beakers were replaced with new units of slightly different geometry. At the same time, spectrum analysis and Cs-137 activity determination transitioned from the in-house Greina software to the GammaVision package.

The present evaluation focused on assessing the performance of these analysis tools across a range of environmental samples, including milk, milk powder, meat, and seawater. Statistical comparison of result parameters showed a strong linear correlation between GammaVision and Greina, with minor divergences attributable to differences in efficiency coefficients across detectors and container geometries. When these differences are corrected, excellent agreement is achieved, confirming the robustness of both methods and the reliability of the updated GammaVision workflow for future monitoring.

The impact of container changes was found to be minor for liquid milk, where the small geometrical differences in Marinelli beakers did not significantly affect measured activity concentrations. In contrast, the change from 200 ml to 350 ml containers for milk powder introduced a systematic positive offset of approximately 1.012 Bq/kg. Correcting for this offset is necessary to maintain continuity of long-term datasets.

For seawater, the updated GammaVision-based analysis reproduces the older method quantitatively, with a small systematic increase (~5%) in activity concentrations due to differences in efficiency calibration. A transitional approach, using legacy Cs-134 coefficients and interpolated Cs-137 efficiencies, has been applied for post-2020 seawater samples to ensure continuity until new, quality-assured calibration measurements are completed.

Finally, IRSA will assess matrix-dependent self-attenuation effects in a planned calibration programme, which will include representative sample matrices and container geometries. This initiative will strengthen the accuracy and traceability of efficiency coefficients, support potential matrix-specific corrections, and ensure continued quality assurance of reported activity concentrations.

Overall, the GammaVision workflow provides a more reliable, flexible, and transparent framework for the determination of Cs-137 in environmental samples, supporting robust long-term environmental monitoring and reporting within the IRSA programme.

1. Introduction

The Icelandic Radiation Safety Authority (IRSA) is tasked with the responsibility of performing monitoring of radiation and radioactivity in the environment. The monitoring program has been maintained since 1986, focusing on measurements of Cs-137 activity concentration in environmental samples, and results in an annual report open to the public [1, 2]. Environmental monitoring in the context of radiation requires a systematic approach to collection and analysis of relevant environmental media, which in Iceland includes milk, milk powder and lamb meat.

In Figure 1, the time evolution of the yearly average activity concentration of Cs-137 in the respective sample type is shown. The figure is reproduced based on data reported in [2].

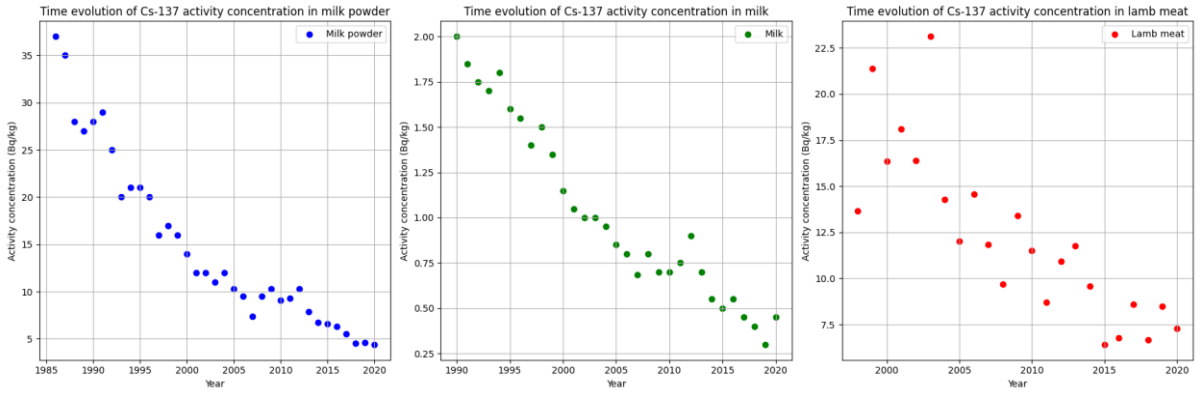


Figure 1: Time evolution of yearly average of Cs-137 activity concentration in milk powder (left subplot), milk (middle subplot) and lamb meat (right subplot), respectively.

Until the year 2021, the Greina software [3] was employed to analyse the gamma spectra obtained from the gamma spectroscopic measurements. The software is limited to quantification of Cs-137 activity only. From 2021, however, the GammaVision software package [4] has been employed to perform the analysis of the gamma spectra to quantify activity concentrations in the measured samples. It should be noted that GammaVision was also prior to 2021 employed as the software for spectrum acquisition, it was merely the analysis step that was performed with Greina by IRSA.

GammaVision is a well-established software package for gamma ray analysis [5] that offers more features than the Greina software, which was the main reason for adopting it. In addition, IRSA started using new sample containers for which Greina was not calibrated: 200 ml containers for media such as milk powder were upgraded to 350 ml, and the 1000 ml Marinelli beakers were exchanged for new units with slightly different geometry. The impact of this change on the resulting analysis is addressed in this analysis.

The present study investigates the quantitative differences between the results for Cs-137 activity concentrations obtained with the two software packages. The performance of each package is compared for identical samples, where large datasets of recorded spectra between 2013 and 2020 were analyzed and the Cs-137 activity quantified.

In addition, upon the change in measurement containers in 2021, the resulting Cs-137 activities for datasets of milk and milk powder are compared pre- and post-2021, using datasets of recorded spectra between 2006 and 2024. In Appendix, table A.1 the different containers and associated volumes that were used for different media pre and post 2021 are shown.

2. Methods

2.1 Data

Gamma spectroscopy measurements of lamb meat, milk and milk powder from the period 2013 to 2020, were analysed using the software packages Greina and GammaVision to assess the agreement between the results for Cs-137 activity concentration obtained using Greina and GammaVision.

For the software package performance assessment, the amount of data points (measured samples) differed depending on the sample type; the milk dataset contained 142 spectrum files, the milk powder dataset 55 spectrum files, and the lamb meat dataset 113 spectrum files, in turn. The Cs-137 activity concentration results obtained with Greina has been previously documented and until 2015, published in yearly reports [1, 2].

For the comparison of container change implication on Cs-137 activity concentration estimation, datasets of milk and milk powder was analyzed. Pre-2021, milk samples were typically analyzed in 1000 ml Marinelli beakers of a certain geometry and post-2021, Marinelli containers of the same nominal volume were employed but with slightly different geometry. For milk powder, both the container type that had been applied previously as well as the nominal volume changed in 2021, from 200 ml to 350 ml. The dataset consisted of monthly measurements of Cs-137 activity concentration in fresh milk from Selfoss (194 samples), fresh milk from Akureyri (199 samples), and milk powder (190 samples) covering the period 2006–2024.

For the seawater analysis, a dataset of 21 seawater samples were employed, for the period 2021-2022.

The measurements were performed using two Ortec HPGe gamma detectors, denoted G2 and G3, at the IRSA research facilities. The datasets were chosen based on data availability and internal documentation.

2.2 High through-put analysis

An automation scheme was developed to facilitate the analyses of the large amount of spectrum data. GammaVision includes a command line interface (CLI) that offers automation up to a certain level, with the utilization of so-called JOB files. The analysis step using GammaVision was performed using the ENV32 analysis engine [4]. Furthermore, an inhouse Python code was developed to fully automate the analysis process.

2.3 Statistical analysis

2.3.1 Performance comparison

The correlation between results obtained with Greina and GammaVision was investigated by means of the Pearson correlation coefficient [6]. This method analyses the linear correlation between two datasets. The correlation coefficient (r-value) measures the strength and direction of a linear relationship between two variables while the p-value indicates the significance, *i.e.* the likelihood that the observed correlation is due to randomness. Linear regression was also performed on the datasets where coefficient of determination (R^2) was used to quantify the linear correlation between the methods. In addition, the root mean square error was computed.

2.3.2 Container and geometry change

To isolate the effect of the 2021 method change from the long-term natural declining trend in Cs-137 concentrations, an ordinary least squares (OLS) regression model was fitted for each sample type according to:

$$\log(y_i) = \beta_0 + \beta_1 t_{index,i} + \beta_2 method_j + \varepsilon_i \,, \qquad (1)$$

where y_i is the Cs-137 activity, $t_{index,i}$ is the sequential time index that captures the long-term declining trend of Cs-137 activity due to radioactive decay and other aspects. The parameter $method_j$ is the binary indicator for container type, which equals 0 or 1 depending on whether the sample is measured pre- or post-2021. This is consistent with approaches used to model step changes in environmental data following procedural interventions [7, 8]. β_0 is the expected log activity at time = 0 using the old container in the data set analyzed. β_1 is the trend parameter and β_2 gives the method effect, i.e. captures the magnitude of the shift in log(y) at the time of the container change. If the analysis results in the conclusion of a significant impact of method change on the Cs-137, the β_2 parameter gives the correction factor. ε is the error term representing the randomness in the measurement that is not explained by time or method.

OLS provides an estimate of how much of the observed change in concentrations can be explained by the passage of time (β_1) versus by the 2021 method change (β_2), with uncertainty quantified through standard errors and p-values.

Although Cs-137 decays exponentially over long periods, the decline over the 18-year study window is sufficiently smooth that a linear time covariate adequately controls for the underlying trend. Ordinary least squares (OLS) regression with a log-transformed response and an indicator variable for the 2021 container change is therefore an appropriate and statistically well-supported method for estimating the method bias [9, 10].

2.4 Determination of Cs-137 activity in seawater

Collected seawater samples are typically ~180 L and thus too large to be measured directly in standard 10 cm diameter vessels and a process needs to be applied to chemically alter the samples prior to gamma spectroscopy measurements; Cs-137 is chemically precipitated and

subsequently measured in the resulting solid fraction. The process includes addition of copper nitrate $Cu(NO_3)_2$ and potassium ferrocyanide $K_4Fe(CN)_6$, respectively, along with a known quantity of Cs-134 as a tracer. These reagents capture both cesium isotopes, which then precipitate over the course of several days. After precipitation, the supernatant was removed, and the resulting solid fraction was analyzed for cesium isotopes. Since isotopes of the same element are chemically indistinguishable, *i.e.* the ion exchange process that occurs in this case is indifferent to the cesium isotope in the solution, the original Cs-137 concentration is calculated from the measured activities of Cs-134 and Cs-137 in the precipitate and the known Cs-134 spike activity. This procedure typically yields 20–50 mL of precipitate which can be measured in specific containers used by IRSA.

The analysis method that is employed to calculate Cs-137 concentration in seawater samples is outlined in Figure 2 below. Measurement data, detector type, and net peak areas for Cs-137 and Cs-134 are extracted from measurement files. Sample information, including seawater mass, processed volume, spike mass, and spike activity, is then retrieved from the laboratory log. Detector- and geometry-specific efficiencies are applied to convert peak areas into activities at the time of counting. The Cs-134 spike activity is decay-corrected to the measurement date and used to calculate chemical yield. Finally, Cs-137 activity is yield-corrected and normalized to the sample mass to obtain the reported seawater activity (Bq/ton).

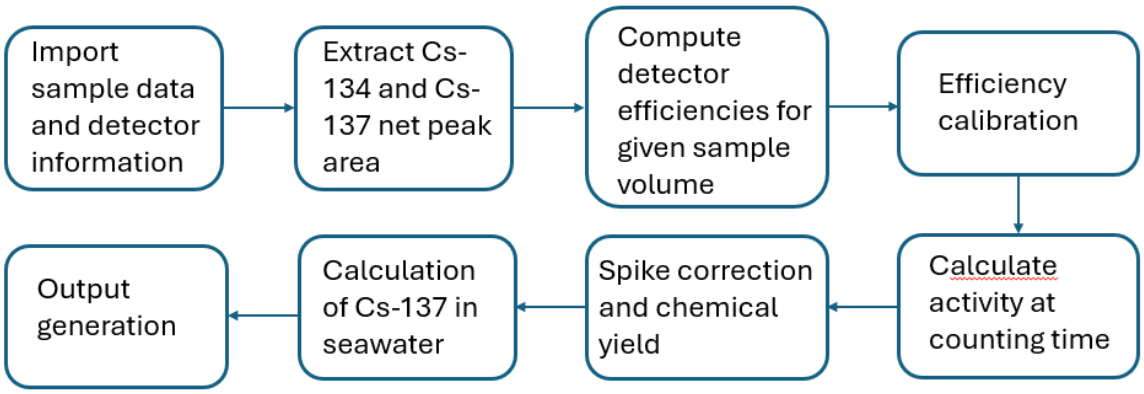


Figure 2: Calculation workflow for determining Cs-137 activity in seawater using a Cs-134 spike. The workflow outlines the full processing sequence used to transform gamma-spectrometric measurements into estimated seawater activity concentrations

Gamma spectra from seawater samples were acquired using HPGe detectors. For each sample, two nuclides were evaluated: Cs-137, representing the analyte of interest, and Cs-134, which was added as a calibrated spike to enable chemical-yield and decay corrections. Detector metadata and measurement parameters (geometry, live time, efficiency calibration) were imported automatically as part of the processing workflow.

Net peak areas for Cs-137 and Cs-134 were obtained from the processed gamma spectra. These peak areas, together with energy-specific efficiency coefficients, were used to convert counts into activities. Efficiency values were taken from calibration curves appropriate for the detector and sample geometry. Activities at the time of counting were calculated using

standard gamma-spectrometric equations, were emission probabilities for Cs-134 and Cs-137 are obtained from Laboratoire National Henri Becquerel [11]. This step yields preliminary activities for both Cs-137 and the Cs-134 spike.

The known activity of the Cs-134 spike at the time of spiking was decay-corrected to the counting date. Comparing the measured Cs-134 activity to the expected activity allowed estimation of chemical yield, accounting for losses during processing:

$$Y = \frac{A_{meas,134}}{A_{true,134}} \quad (2)$$

This yield factor was used to correct the Cs-137 activity accordingly. The Cs-137 activity recovered from the spectrum was divided by the chemical yield and normalized to the processed seawater volume to obtain the final seawater activity concentration:

$$A_{Cs-137,seawater} = \frac{A_{Cs-137,meas}}{Y \cdot V_{sample}} \quad (3)$$

Which gives the results in Bq/ton, consistent with IRSA reporting standards on Cs-137 activity in Icelandic seawater. The workflow described above has been implemented into a Python code for automated analysis.

The change in analysis approach on data obtained after 2021 differ in the step where the detector efficiency is calculated, due to different efficiency coefficients use in Greina, which is used in the older approach, and GammaVision, which is used in the new method, and an update in the sample volume dependent interpolation scheme applied in the calculation.

During the comparison of GammaVision and Greina performance, see section 3.1, it was concluded that the calibration performed for the 200 ml DOS container was faulty for higher volumes (section 3.1.1). Hence, a temporary a solution was adapted, pending a new calibration measurement, where the efficiency coefficients from the old calibration, employed in Greina, is employed for Cs-134. For Cs-137, an interpolation power-law algorithm using reliable efficiency coefficients for lower volumes for the 200 ml DOS container is adapted until new calibration measurements are performed.

In addition, updated emission probability constants are introduced due to development in the field since the original analysis code was written. This update has insignificant effects, however.

The difference in results with the new and old analysis tool is compared in the present study, see section 3.3. To assess agreement between the methods, the differences between the results from the two methods, the Greina based and the GammaVision based, and their means were calculated. Descriptive statistics, including mean, standard deviation, and range of differences, were used to summarize systematic bias. Statistical significance of the differences was evaluated using a paired t-test, with a Wilcoxon signed-rank test applied to account for potential non-normality. Pearson correlation coefficients were calculated to assess the linear association between the two methods. Agreement was further examined using a Bland–Altman plot [8], with the mean difference and 95% limits of agreement (mean ± 1.96 × standard deviation of differences) visualizing systematic and random variation. A

linear regression of the GammaVision based results and those from the older Greina method was also performed to illustrate the overall consistency between the methods.

3. Results

3.1 Comparison of GammaVision and Greina performance

Activity concentrations of Cs-137 as calculated with GammaVision were plotted against previously documented data for the same parameter as obtained with Greina. Figure 3 shows three 1:1 ratio scatter plots of the three datasets respectively, plotted along with a 1:1 reference line.

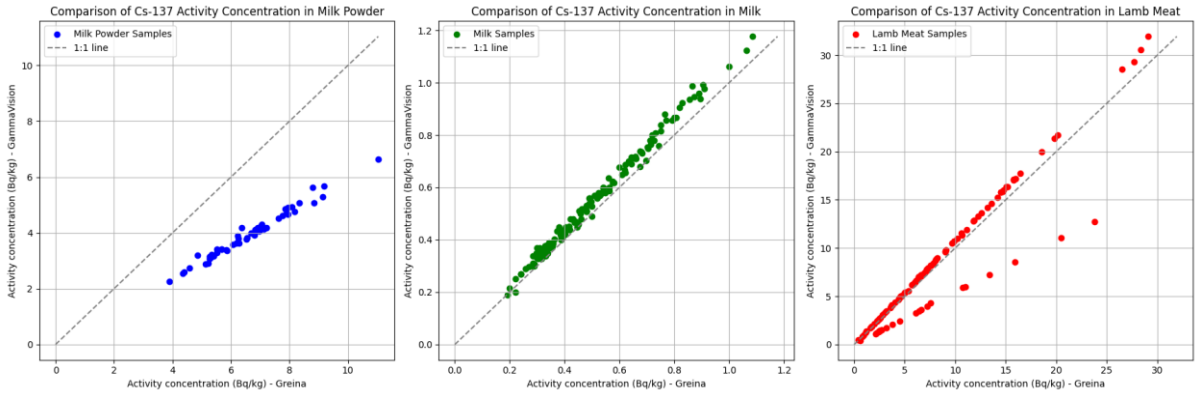


Figure 3: Left: 1:1 ratio scatter plot comparing Greina Cs-137 activity concentrations (x-axis) with GammaVision Cs-137 activity concentrations (y-axis) for the milk powder data set. Center: 1:1 ratio scatter plot comparing Greina Cs-137 activity concentrations (x-axis) with GammaVision Cs-137 activity concentrations (y-axis) for the milk data set. Right: 1:1 ratio scatter plot comparing Greina Cs-137 activity concentrations (x-axis) with GammaVision Cs-137 activity concentrations (y-axis) for the lamb meat data set.

Starting with the milk powder comparison, it is noted that the Greina analysis typically yields higher activity concentrations, and the discrepancy appears to diverge linearly with increasing activity concentration. For the milk data set, there seems to be a reasonable agreement between the two approaches. For the lamb meat data set, on the other hand, two different behaviours can be observed. One subgroup of the data appears to diverge linearly, similarly to the observation for milk powder, where the Greina analyses have yielded higher activity concentrations. Another subgroup seems to compare better, albeit still with a noticeable linear divergence with increasing activity concentration; for this group however the GammaVision results are typically higher than those from Greina calculations.

To further assess the comparison between the GammaVision and Greina results, the correlation between the GammaVision and Greina results were analyzed, along with root mean square error. The results from the analysis are displayed in Table 1.

Table 1: Root mean square error (RMSE) and Pearson correlation coefficient (p- and r-values) for the milk powder, milk and lamb meat datasets.

Dataset	RMSE	Pearson r	Pearson p-value
Milk Powder	2.73	0.99	8E-46

Milk	0.050	0.99	8E-151
Lamb Meat	2.06	0.95	3E-58

The RMSE values indicate varying levels of absolute agreement, which is also visually observable with the divergence earlier mentioned. The linear correlation is however strong (near 1) for all datasets, and the significance is statistically significant (< 1E-45) [6].

To further understand the nature of the linear divergence the relative bias is calculated, see Figure 4. This was performed to assess whether observed behaviour is systematic, i.e. whether GammaVision systematically under- or overestimated the activity concentration within the different datasets. Relative bias was for this analysis calculated according to:

$$Relative\ bias\ (\%) = \frac{A_{GV} - A_{Greina}}{A_{Greina}} \times 100. \qquad (4)$$

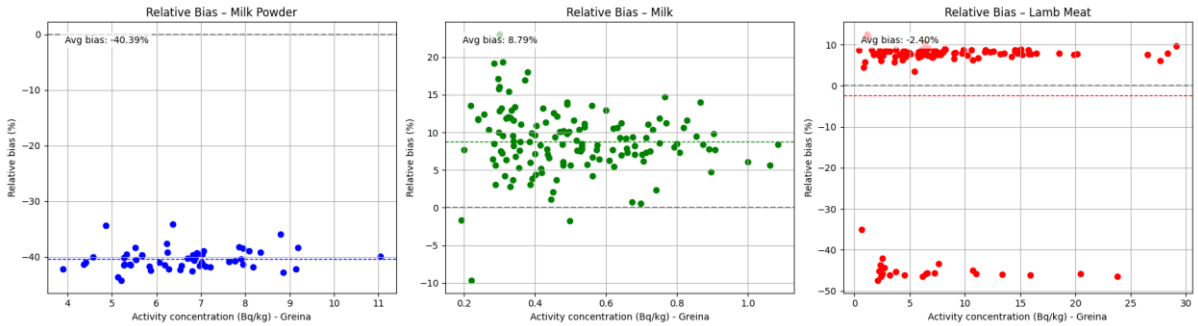


Figure 4: Relative bias calculated for the milk powder (left), milk (center), and lamb meat (right) datasets.

The relative bias for milk powder clearly shows that GammaVision is consistently yielding activity concentrations that are around 40% smaller than that obtained with Greina. For milk, the average bias is about 9%. For lamb meat the calculated average bias is -3% but this does not reflect the observation that there are at least two subgroups in the dataset with consistent over- and underestimation of the activity concentration calculated with GammaVision compared to that computed with Greina. Grouping the lamb meat data according to the different sample volumes employed presents a more comprehensive analysis of the comparison, see Figure 4.

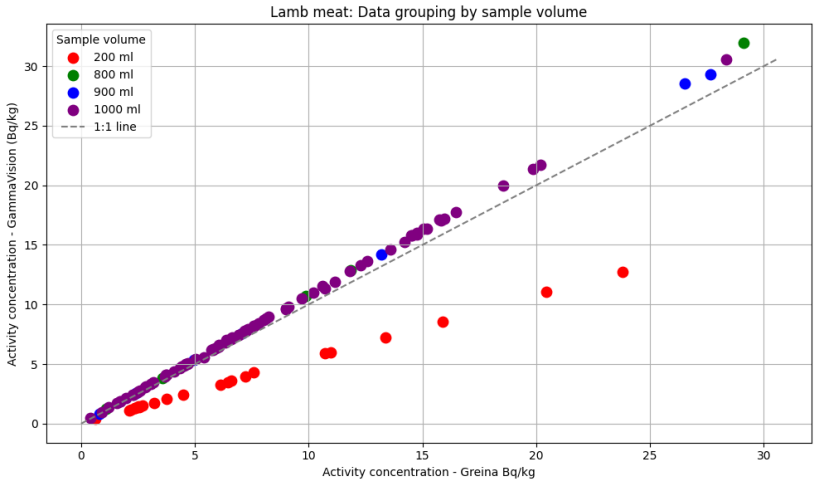


Figure 5: Comparison of activity concentration of Cs-137 in lamb meat obtained with Greina and GammaVision partitioned into subgroups depending on sample volume employed in the gamma spectroscopic measurement (200 ml (red), 800 ml (green), 900 ml (blue) or 1000 ml (purple)).

Figure 5 demonstrates that the linear divergences are connected to the sample volume employed during the measurements. The discrepancy appears to be largest for the 200 ml subgroup (red dots in Figure 5), while the other sample volumes seem to follow the same linear behaviour.

The reasons for this systematic discrepancy could be manifold. However, a common reason for systematic errors or biases can be due to differences in efficiency calibration. This has been discussed in-depth in textbook literature [12, 13].

Different sample geometries and volumes were used both between and within the datasets for the gamma spectrometry measurements and different efficiencies were employed for the respective software. The employed efficiencies also depend on the gamma detector used for the measurement. IRSA has two stationary gamma detectors installed, denoted G2 and G3. Table 2 shows the efficiency values employed for the two programs for the energy 661.6 keV (Cs-137).

Table 2: Efficiency calibration values employed for GammaVision (GV) and Greina for different measurement configurations (sample geometries and volumes. MAR denotes Marinelli beaker and DOS denotes a smaller plastic geometry employed previously by the IRSA. The efficiencies used by the respective programs are given for the energy 661.6 keV (Cs-137).

Measurement configuration			**Efficiency (Greina)**	**Efficiency (GV)**	**Percentage difference (Greina vs. GV) (%)**
Detector	**Geometry**	**Volume (ml)**			
G2	MAR	1000	0,021	0,019	+7,8
G3	MAR	1000	0,017	0,015	+7,9
G2	MAR	900	0,022	0,020	+8,2
G3	MAR	900	0,018	0,016	+1,2
G2	MAR	800	0,023	0,021	+7,0
G3	MAR	800	0,019	0,017	+9,6

G2	DOS	200	0,019	0,034	-43,4
G3	DOS	200	0,018	0,034	-46,1

As seen in Table 2, for most of the geometries employed the efficiency value differs. In general, the employed Greina efficiencies are in the order of 8% higher for all the employed Marinelli geometries while for the DOS geometry, the GammaVision efficiency used is about 40% larger than that of Greina. The reason for this large discrepancy is analyzed further in section 3.1.1 below.

A general expression for the calculation of activity in gamma spectrometry is given in equation 2 below:

$$A = \frac{N/t_g}{\varepsilon \times P_E \times q}, \qquad (5)$$

where A is the activity, N is the net count rate, ε is the full energy peak efficiency at energy E, P_E is the gamma emission probability, t_g is the measurement live time, and q, the sample quantity in e.g. kg or m^3.

To correct for the different efficiency values, we define, based on expression 2, an efficiency correction factor that translates the GammaVision activity concentration to quantitively compare with the Greina activity concentrations for each measurement configuration *i*:

$$A_{GV\_corr,\ i} = \frac{\varepsilon_{GV,i}}{\varepsilon_{Greina,i}} \times A_{GV} = \varepsilon_{corr,i} \times A_{GV}. \qquad (6)$$

The configuration *i* is defined based on the measurement variables volume, geometry and detector, see Table 2.

The efficiency correction factors for the different measurement configurations that quantifies the discrepancy between calculated Greina and GammaVision activities are shown in Table 3.

Table 3: Efficiency correction factors, $\varepsilon_{corr,i}$, calculated for the measurement configurations employed for the samples included in the milk powder, milk and lamb meat datasets investigated in the present study.

Measurement configuration			
Detector	**Geometry**	**Volume (ml)**	**$\varepsilon_{corr,i}$**
G2	MAR	1000	0.90
G3	MAR	1000	0.88
G2	MAR	900	0.91
G3	MAR	900	0.89
G2	MAR	800	0.91
G3	MAR	800	0.89
G2	DOS	200	1.79
G3	DOS	200	1.89

Indeed, applying the efficiency correction factors on the lamb meat data set shows that the origin of the observed linear divergence with increasing activity concentration, as can be seen in Figure 6. Now, the efficiency corrected activity concentrations follow a similar linear behaviour regardless of sample volume employed in the measurements.

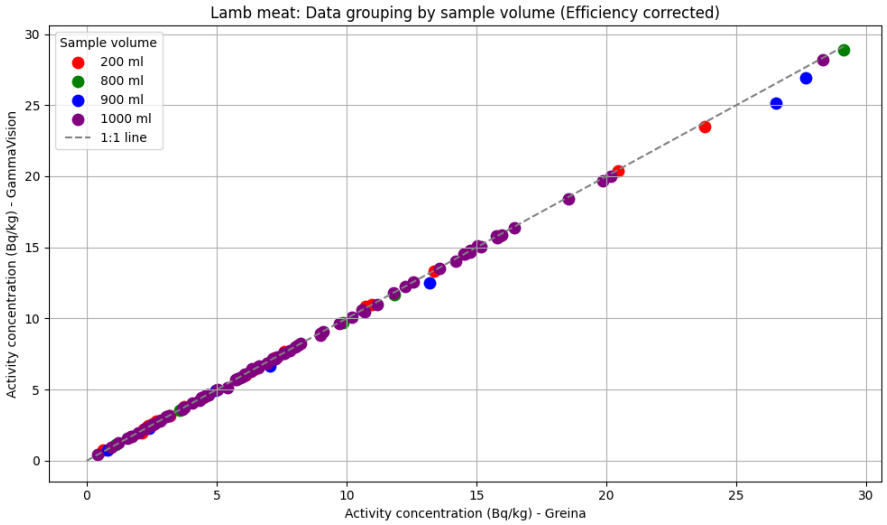


Figure 6: Efficiency corrected activity concentrations for the lamb meat dataset. The same partitioning and colours are employed as for Figure 3 (200 ml (red), 800 ml (green), 900 ml (blue) or 1000 ml (purple)).

Efficiency corrected activity concentrations for all three datasets are shown in Figure 7 along with linear regression plots to analyse the comparison.

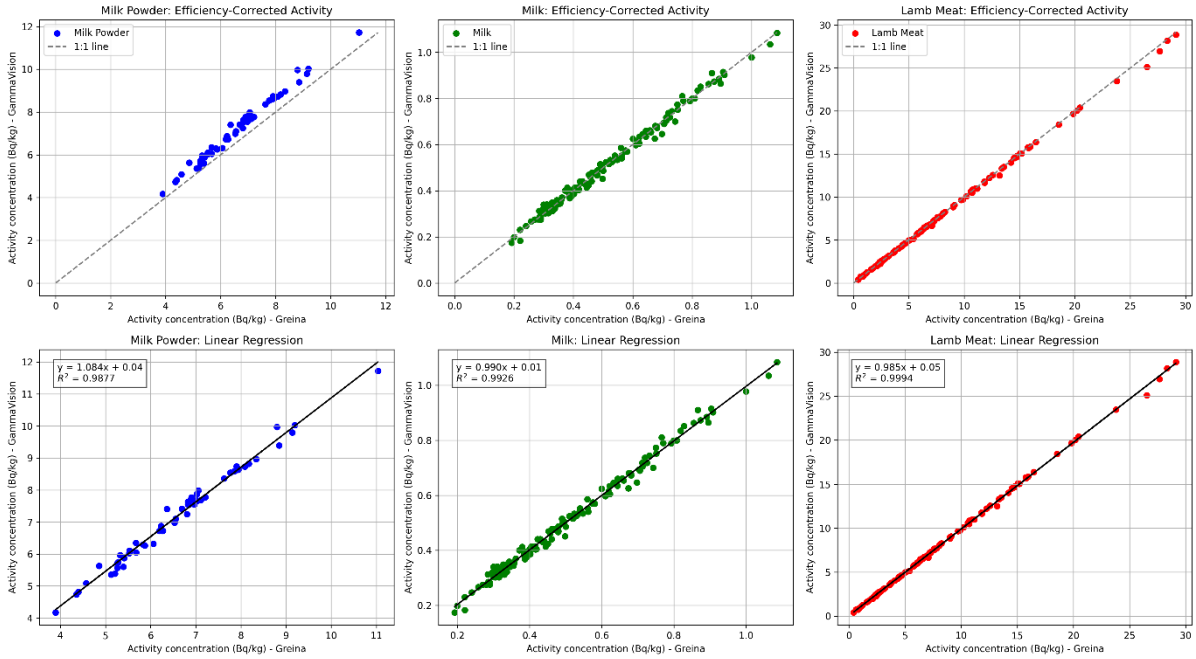


Figure 7: Efficiency corrected activities (according to expression 3) for milk powder (top left), milk (top center) and lamb meat (top right). Including linear regression: milk powder (bottom left), milk (bottom center), lamb meat (bottom right).

As can be seen in Figure 7, the correction made yields good agreement between the two methods, as is confirmed by the coefficient of determination (R^2), which is 0.99 for all datasets. Thus, the linear divergence observed in the initial comparison can primarily be attributed to the efficiency calibration employed in the different software codes.

Additional discrepancies between the two methods' approaches to quantify activity may exist, e.g. in how they treat peak fitting, background subtraction and decay correction handling. However, that is beyond the scope of this study to assess.

The derived correction factors displayed in Table 3 for the Marinelli containers will be considered in continued reporting of activity concentration within the IRSA environmental monitoring program to illustrate the observed discrepancies between the two software approaches. The large discrepancy observed for the 200 ml containers is discussed further in section 3.1.1.

3.1.1 Difference in efficiency coefficients for 200 ml containers

Due to the large discrepancy observed in the efficiency coefficient for the 200 mL DOS container (Table 2), the efficiency calibration performed by IRSA in 2020 was revisited. Table 4 presents the efficiency coefficients for Cs-137 obtained for different sample volumes measured in the 200 mL DOS container using detectors G2 and G3.

Table 4. Efficiency coefficients for Cs-137 (661.66 keV) for different sample volumes measured in 200 ml DOS containers. Values are shown for the detectors G2 and G3, respectively.

	Efficiency coefficients for Cs-137 (661.7 keV)	
Volume	G2	G3
30	3.62E-02	3.55E-02
35	3.55E-02	3.47E-02
45	3.42E-02	3.28E-02
50	3.21E-02	3.15E-02
80	2.77E-02	2.70E-02
100	2.52E-02	2.48E-02
200	3.57E-02	3.41E-02

The efficiency value reported for the 200 mL volume does not follow the expected trend of decreasing efficiency with increasing sample volume, suggesting a likely error in the original sample preparation. To investigate this further, the calibration solution, which has been stored at IRSA facilities since the 2020 calibration procedure, was remeasured in an alternative container with a reliable calibration. This remeasurement confirmed that the activity of the solution was approximately twice as high as expected based on the documentation from the 2020 calibration procedure, consistent with the anomalously high efficiency coefficient obtained for the 200 mL volume.

Based on these findings, it is deemed necessary to revert to the efficiency calibration coefficients determined for the DOS geometries prior to 2020 for gamma spectrometric analysis, until a new calibration procedure can be performed.

3.2 Impact of container change on estimated Cs-137 activity concentrations

OLS regression results for each sample dataset are summarized in Table 5. The binary method variable (β_2) represents the estimated shift associated with the 2021 container change after adjusting for the long-term decline.

Table 5. OLS regression estimates of method effect (β_2), 95% confidence intervals (CI), p-values, and sample sizes (N) for the three datasets (Milk Selfoss, Milk Akureyri and Milk powder).

Sample type	β_2 - method effect (Bq/kg)	95% CI	p-value	N
Milk (Selfoss)	0.056	-0.030 to 0.14	0.20	194
Milk (Akureyri)	0.026	-0.012 to 0.064	0.18	199
Milk powder	1.012	0.296 to 1.73	0.0058	190

As discussed in section 2.3.2, the parameter β_2, which indicates the "method effect", captures the magnitude of the shift in Cs-137 activity concentration at the time of the container change. The estimated method effect differed across the datasets where no significant shift was observed for either of the milk datasets, but a significant upward shift (1.012 Bq/kg) was observed for the milk powder dataset.

Statistical significance was assessed using the conventional $\alpha = 0.05$ level, meaning that p-values below 0.05 were interpreted as evidence of a non-zero method effect, which is the case for the milk powder dataset. P-values above this threshold were considered insufficient to conclude that the observed differences reflect anything other than random sampling variability [14], which is the case for the milk datasets, see table 5.

A 95% confidence interval (CI) defines the range of values within which the true method effect is expected to lie with 95% probability, assuming the model is correct. If the CI does not include zero, which is the case for milk powder, this indicates that the estimated effect is statistically distinguishable from no effect at conventional significance levels ($\alpha = 0.05$). Conversely, if the CI does include zero, we cannot rule out that the apparent difference between pre- and post-2021 measurements is due simply to random variation rather than a systematic analytical shift.

Figure 8 illustrates these observations for the three datasets. As can be seen in the figure, there is a noticeable decline in Cs-137 activity for all datasets, as expected. For the milk datasets, top and middle plots, there are visible upshifts at the onset of 2021, when the container change was implemented.

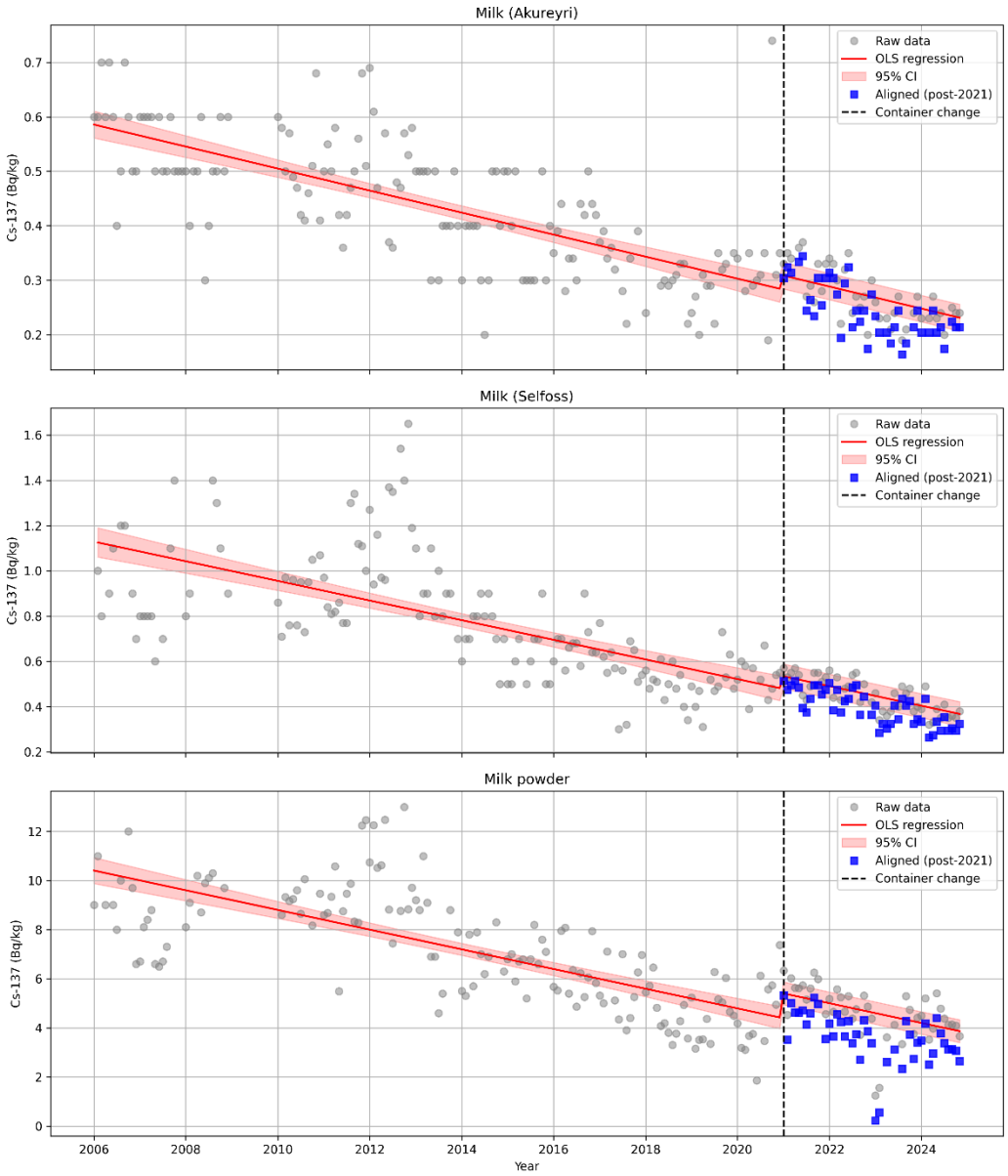


Figure 8: OLS regression analysis of the three datasets, Milk (Akureyri), Milk (Selfoss) and Milk powder. The top subplot shows raw and post-2021 aligned activities for milk from Akureyri, the middle subplot shows milk from Selfoss, and the bottom subplot shows milk powder. Gray circles represent measured data, red lines indicate ordinary least squares (OLS) regression fits, and shaded red areas correspond to 95% confidence intervals for the regression. Blue squares show post-2021 aligned values after adjusting for potential changes in measurement method. The vertical dashed line marks the container/method change in 2021.

As shown in Table 5, these are however not significant, and post-2021 alignment (blue squares in the figure) shows only minimal adjustments. In contrast, milk powder exhibits a small but noticeable shift after 2021, which is captured by the alignment procedure. The regression analysis and confidence intervals suggest that the observed trends in milk powder likely reflect methodological changes rather than environmental variation. Overall, these results demonstrate that milk Cs-137 levels remained largely consistent, whereas adjustments are necessary for milk powder to account for container-related measurement differences after 2021.

The analysis pre- and post-2021 is, as already mentioned, performed with Greina and GammaVision, respectively. The difference in efficiency calibration applied in the respective software for the old Marinelli geometry indicates that an efficiency correction factor of about 0.9 depending on detector, see Table 3, is needed for a quantitative comparison of GammaVision results with activity concentrations obtained with Greina. Since the post-2021 measurements are performed with the new Marinelli geometry, it is not possible to perform a quantitative analysis with Greina on that subgroup of the dataset. The efficiency coefficients for the Cs-137 energy obtained from the calibration measurements and applied in the analysis of the post-2021 data is similar for the old and new Marinelli geometries 0.016 and 0.19, for the G3 detector, and 0.019 in both cases for the G2 detector. Thus, a similar correction factor of 0.9 can be qualitatively applied to assess the impact of difference in efficiency calibration on the comparison of activities pre- and post-2021, see Figures 9 and 10 for the Milk Akureyri and Milk Selfoss datasets respectively.

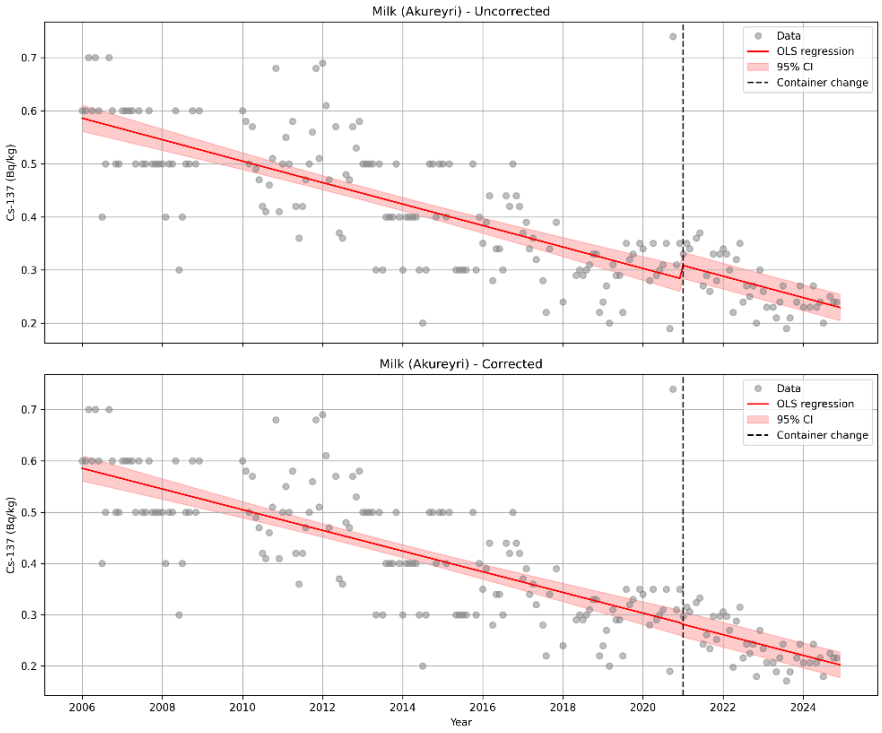


Figure 9: Impact on regression analysis without (top) and with (bottom) an efficiency correction factor of 0.9 applied on the Milk Akureyri dataset.

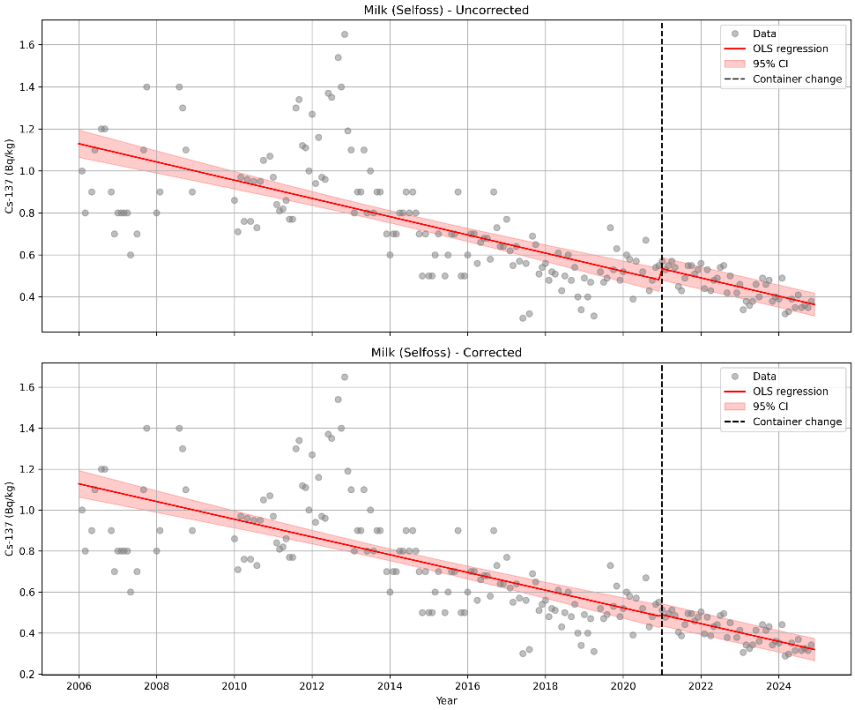


Figure 10: Impact on regression analysis without (top) and with (bottom) an efficiency correction factor of 0.9 applied on the Milk Selfoss dataset.

As can be seen in Figures 9 and 10, qualitative consideration of efficiency calibration renders excellent agreement for the milk datasets. This exercise was however merely to illustrate the effect rather than to correct the data, since no such operation is deemed necessary on the milk datasets applying containers of similar geometry, based on the results from the statistical analysis shown in Table 5 and Figure 8.

In summary, after accounting for the long-term decline, the analytical change in 2021 had no statistically significant effect on the two liquid milk datasets. However, results for milk powder increased by approximately +1.01 Bq/kg relative to the pre-2021 regime, indicating a measurable positive method bias due to the change in container volume and geometry for the milk powder samples.

3.3 Updated approach to quantification of Cs-137 in seawater

Cs-137 activity measured in seawater using the GammaVision and Greina based methods showed a strong linear association (Pearson $r = 0.95$, $p < 0.001$), indicating that the two methods co-vary closely across the observed activity concentration range. However, paired comparisons revealed a small but statistically significant difference between methods see Table 6, with GammaVision based results tending to be slightly higher than results based on Greina.

Table 6: Statistical comparison of Cs-137 activity measurements in seawater using the GammaVision and Greina based methods. The table reports the mean difference (GammaVision – Greina), standard deviation of differences, 95% limits of agreement, results of the paired t-test, and the Pearson correlation coefficient. While GammaVision based results are slightly higher on average than GREINA (mean difference = 0.045 Bq/ton), the two methods show a strong linear association (r = 0.95), and the limits of agreement indicate that 95% of differences fall between –0.073 and 0.163 Bq/ton.

Statistic	Value
Mean difference	0.045 Bq/ton
Standard deviation of difference	0.060 Bq/ton
95% limits of agreement	-0.073 to 0.16 Bq/ton
Paired t-test	t = 3.43, p = 0.0026
Pearson correlation	r = 0.95, p < 0.001

The agreement between the methods is further visualized in Figure 11.

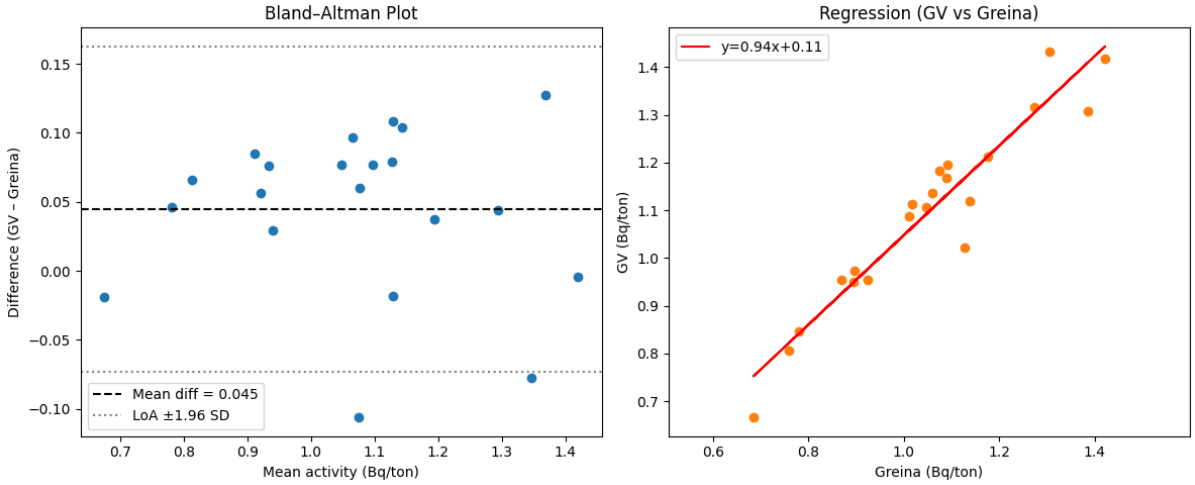


Figure 11: Comparison of Cs-137 activity concentration in seawater using GammaVision and Greina based methods. Left: Bland–Altman plot showing the differences between GV and GREINA measurements (GV – GREINA) plotted against their mean values. The dashed line represents the mean difference (bias), and the dotted lines indicate the 95% limits of agreement (mean ± 1.96 × SD), illustrating the extent of agreement and variability between methods. Right: Scatter plot with linear regression of GammaVision versus Greina based Cs-137 concentrations, with the regression line indicating the overall relationship.

The Bland–Altman plot (left subplot) shows that 95% of differences fall within –0.073 to 0.163 Bq/ton, suggesting reasonable overall agreement despite the small systematic bias. The regression plot (right subplot) further illustrates the strong linear relationship between the two methods, with the regression line closely approximating the identity line.

No trend in the differences was observed across the measurement range, indicating consistent bias. The mean paired difference was 0.045 ± 0.060 Bq/ton, corresponding to about 4.5%. The difference was statistically significant (paired t-test: t = 3.43, p = 0.0026) but despite the significance, the magnitude of the bias is minor relative to analytical uncertainties and environmental variation. Although statistically detectable, the magnitude of the difference is small relative to typical Cs-137 variability in seawater and therefore unlikely to have practical significance. Nonetheless, Cs-137 concentrations obtained with the new method is expected to be on average about 5% higher than the old method.

4. Discussion

4.1 Impact of methodological changes within the IRSA monitoring program

A systematic comparison of the performance of the Greina and GammaVision software packages was conducted as part of the evaluation of methodological changes in the IRSA monitoring programme. In general, the results obtained using the two approaches show good agreement. Minor differences were observed and can primarily be attributed to changes in the calculation workflow and implementation details associated with the transition between software packages.

In a few cases, larger discrepancies between the two methods were identified. Further investigation demonstrated that these discrepancies originate from differences in the applied efficiency calibrations rather than from the spectral analysis itself. When consistent efficiency coefficients are applied, the results from Greina and GammaVision show essentially perfect agreement, confirming the robustness of the underlying activity determination methodology.

The impact of changes in sample containers implemented in 2021 was evaluated using datasets of milk and milk powder. For liquid milk, Marinelli beakers with identical nominal volumes but slightly different geometries were used for measurements conducted before and after 2021. For milk powder, samples were measured in 200 ml containers prior to 2021, while a larger 350 ml container was introduced thereafter.

For the milk datasets, the comparison revealed a detectable but statistically insignificant effect of the container change on the derived activity concentrations. On this basis, no correction for container-related differences is considered necessary when comparing results obtained before and after 2021. This conclusion is expected to be applicable to other sample media with similar physical properties and measurement geometries, including lamb meat, fish, and seaweed, although these media were not explicitly included in the present evaluation due to the limited number of available data points.

In contrast, for milk powder, for which the measurement container was changed from 200 ml to 350 ml in 2021, a systematic difference in the derived activity concentrations was identified. When comparing data obtained before and after 2021, an expected positive offset of approximately 1.012 Bq/kg on data post-2021 should therefore be considered or explicitly reported, in order to ensure consistency across the time series.

A more detailed examination of the efficiency calibration data revealed faulty calibration coefficients associated with the DOS container at a volume of 200 ml. Based on this finding, it was decided to undertake a new set of rigorous and quality-assured efficiency calibration measurements for all containers and volumes currently employed by IRSA. This action is intended to ensure the continued reliability and traceability of reported results within the monitoring programme.

Until the new calibration measurements have been completed and validated, a transitional approach will be applied for affected datasets. The previously established efficiency

coefficients will be retained for containers for which the faulty calibration was identified, which specifically applies for seawater samples post-2020. For these samples, legacy efficiency coefficients for Cs-134 are used, while a power-law interpolation based on more recent efficiency data for Cs-137 at lower DOS volumes is applied. This approach enables the continued use of GammaVision within the analysis workflow while maintaining consistency and minimizing systematic bias in the reported activity concentrations.

This approach was used in section 3.3 to compare with the older method which was based on Greina and other in-house developed codes. The comparison identified a small systematic difference between the methods, which can be attributed primarily to differences in the applied efficiency calibrations rather than to the spectral analysis itself. While the magnitude of this difference is small relative to typical analytical uncertainties and the natural variability observed in environmental seawater samples, it should nevertheless be acknowledged in the interpretation of long-term monitoring data. This systematic difference, where results obtained using the updated methodology are expected to be approximately 5% higher than those derived using the earlier approach, should be taken into account when comparing Cs-137 activity concentrations in seawater across the methodological transition in future environmental monitoring reports.

Overall, the transition to the GammaVision-based analysis workflow represents an improvement in the robustness and transparency of the Cs-137 activity determination within the IRSA monitoring programme. GammaVision provides a more flexible and comprehensive framework for efficiency handling and uncertainty evaluation, and the comparative analyses performed demonstrate that, when consistent efficiency calibrations are applied, the results are fully compatible with those obtained using the previous GREINA-based approach. On this basis, results generated using GammaVision are considered to provide a more reliable foundation for future reporting and long-term environmental monitoring.

Nevertheless, the identification of faulty efficiency calibration coefficients for specific container–volume combinations highlights the importance of maintaining a rigorous and traceable calibration framework. Hence, performing new, quality-assured efficiency calibration measurements remains essential to fully validate and consolidate the updated methodology. Until these calibrations have been completed, the transitional approach described above (applied to seawater samples post-2021) ensures continuity and minimizes systematic bias, while ensuring the comparability and integrity of reported results.

4.2 Water-based efficiency calibration for environmental samples

In the IRSA monitoring program, detector efficiency coefficients are derived from calibration measurements performed using aqueous reference solutions. Consequently, activity concentrations in environmental samples are calculated under the assumption that their physical and radiological properties are equivalent to those of water. This assumption has both advantages and limitations, depending on sample composition, density, and geometry [15].

For Cs-137 measurements by gamma spectrometry, the primary consequence of using water-based efficiency calibration for non-aqueous environmental samples is matrix-dependent gamma self-attenuation. This effect arises from differences in density, composition, and effective atomic number between the sample matrix and water.

At the Cs-137 gamma energy (662 keV), photoelectric absorption is negligible, and gamma interactions are dominated by Compton scattering. Consequently, attenuation depends mainly on electron density, which is closely related to bulk density and hydrogen content. Samples with higher density or lower hydrogen content than water will attenuate gamma rays more strongly, resulting in reduced detection efficiency.

There are also secondary effects that may contribute, such as e.g. heterogeneity of radionuclide distribution. In addition, coincidence summing and scattering differences may have effects but for Cs-137 these are negligible [15].

In summary, differences in sample self-attenuation relative to water directly affect the full-energy peak efficiency. If water-based efficiency calibration is applied to denser or drier matrices without correction, the detection efficiency is overestimated, resulting in a systematic underestimation of the calculated activity concentration. The magnitude of this bias depends primarily on sample density and composition [13, 15].

To address and further quantify the impact of matrix-dependent self-attenuation, IRSA will assess these effects as part of a planned programme of new efficiency calibration measurements. This calibration effort will include representative sample matrices and container geometries used in the monitoring programme, with the aim of improving the accuracy and traceability of efficiency coefficients. The results will be used to evaluate the need for matrix-specific corrections or uncertainty contributions and to ensure the continued robustness and quality assurance of reported activity concentrations.

5. Summary and conclusion

5.1 Comparison of Analysis Methods

A systematic evaluation of the GREINA and GammaVision software packages demonstrated good overall agreement in calculated Cs-137 activity concentrations. Minor differences were attributable primarily to differences in efficiency calibration and calculation workflow. When consistent efficiency coefficients are applied, results from both methods are effectively equivalent, with GammaVision providing a more reliable and transparent framework for future analyses.

5.2 Impact of Container Changes

The change in sample containers in 2021 had negligible impact on liquid milk measurements, while for milk powder and other small-volume samples, a systematic offset (~1.012 Bq/kg) was identified. Corrections or explicit reporting of this offset are recommended when comparing pre- and post-2021 data.

5.3 Updated Approach to Quantification of Cs-137 in Seawater

The implementation of GammaVision for seawater analysis provides a more flexible and transparent framework. A small systematic difference (~5%) relative to previous methods was observed due to efficiency calibration, which should be considered in long-term trend analysis. Transitional procedures using legacy Cs-134 coefficients and interpolated Cs-137 efficiencies ensure continuity until new calibrations are completed.

5.4 Matrix Effects and Efficiency Calibration

Water-based efficiency calibrations are currently used for all environmental samples, with self-attenuation in denser or drier matrices being the primary source of potential bias. IRSA plans to perform new, quality-assured calibration measurements to assess matrix-dependent effects and improve the accuracy and traceability of efficiency coefficients across relevant sample types.

5.5 Automation of Analysis Workflow

A high-throughput analysis scheme was developed using the command-line interface (CLI) features of GammaVision and Python scripting. This approach streamlines data processing, ensures consistent application of efficiencies, and supports efficient and reproducible reporting for the IRSA monitoring programme.

References

1. **GR 16:01**, K. Guðnason, R. K. Lárusson, S. Gunnarsdóttir, G. Jónsson, *Radioactivity in the environment and food in Iceland 2015*, Icelandic Radiation Safety Authority, 2016

2. **GR 2025:02**, *The Icelandic Radiation Safety Authority Environmental Monitoring Measurements 2016-2020*, Icelandic Radiation Safety Authority, 2025.

3. **Greina**, Icelandic Radiation Safety Authority.

4. **GammaVision** 8.10, Ortec/Ametek.

5. **IAEA-TECDOC-1011**. *Intercomparison of gamma ray analysis software packages.* International Atomic Energy Agency, 1998.

6. **J. R. Taylor.** *An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements.* 2nd Ed. University Science Books, 1997.

7. **Montgomery, D.C., Jennings, C.L., and Kulahci, M**. *Introduction to Time Series Analysis and Forecasting*. Wiley. 2015.

8. **IAEA**. E*nvironmental Radioactivity: Measurement, Analysis and Interpretation*. Vienna: International Atomic Energy Agency. 2013.

9. **Mück, K.** *Time dependence of the transfer of radiocaesium to cow's milk in Austria after the Chernobyl accident*. Health Physics*, 72(3), 380–390.* 1997.

10. **Gilbert, R. O.** *Statistical Methods for Environmental Pollution Monitoring.* John Wiley & Sons. 1987.

11. http://www.lnhb.fr/home/nuclear-data/nuclear-data-table/

12. **G. R. Gilmore.** *Practical Gamma-Ray Spectrometry*. Wiley. 2008.

13. **G. F. Knoll.** *Radiation Detection and Measurement*. 4th edition. Wiley. 2010.

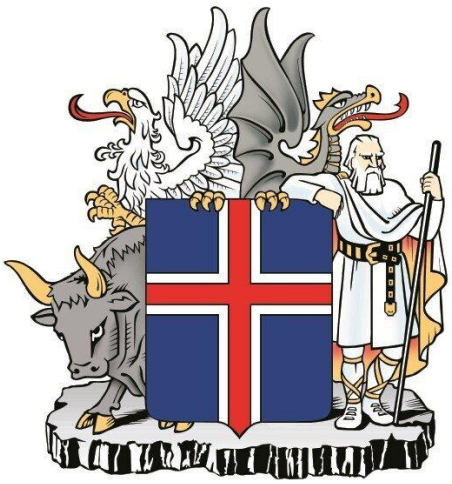
14. **Fisher, R. A.** *Statistical Methods for Research Workers.* Edinburgh: Oliver & Boyd. 1925.

15. **Kaminski, S., Jakobi, A., Wilhelm, C.** *Uncertainty of gamma-ray spectrometry measurement of environmental samples due to uncertainties in matrix composition, density and sample geometry.* Applied Radiation and Isotopes, 94, 306-313. 2014.

6. Appendix

Table A.1: Container types and associated volumes employed by the Icelandic Radiation Safety Authority for different sample types before and after 2021.

	Before 2021		After 2021	
Sample type	**Container type**	**Volume (ml)**	**Container type**	**Volume (ml)**
Milk	Marinelli	1000	Marinelli (new)	1000
Milk powder	DOS (GG)	200	DOS (HV)	350
Lamb meat	Marinelli	500-1000	Marinelli (new)	500-1000
Fish	Marinelli	500-1000	Marinelli (new)	500-1000
Seaweed	Marinelli	1000	Marinelli (new)	1000
Seawater	DOS (GG)	200	DOS (GG)	200
Precipitation	DOS (GG)	200	DOS (GG)	200



GEISLAVARNIR RÍKISINS
ICELANDIC RADIATION SAFETY AUTHORITY