HONG KONG OBSERVATORY

Technical Note (Local) No. 84

Performance of Hong Kong Observatory in Inter-laboratory Comparison Exercises on Radioactivity Measurements 1989 to 2005

by

K.C. Hui, S.W. Li and K.C. Tsui

Published March 2007

Prepared by

Hong Kong Observatory
134A Nathan Road
Kowloon
Hong Kong

551.510.7

摘要

香港天文台自一九八九年開始參與國際實驗室間測量比對活動。這些比對活動有助評估香港天文台輻射實驗室在輻射測量工作中的表現。

此報告總結了天文台由一九八九年至二零零五年在比對活動中取得的成績，並透過不同的指標對天文台輻射實驗室的表現作出評價。分析結果顯示大部份天文台實驗室的測量結果都在比對活動主辦機構公佈的可接受範圍之內。天文台在比對活動中的表現平均是屬於較佳表現的實驗室。

ABSTRACT

The Hong Kong Observatory (HKO) has been participating in international inter-laboratory comparison exercises since 1989. These comparison exercises serve to gauge the performance of HKO's Radiation Laboratory in radioactivity measurements.

This report summarizes the results obtained by the Observatory in the comparison exercises from 1989 to 2005 and assesses the performance of its Radiation Laboratory through various indicators. It was found that a majority of the measurement results of HKO's laboratory fell within the acceptable ranges announced by the exercise organizers. The performance of HKO on average ranked among the group of better performing laboratories in the exercises.

CONTENTS

		page
	ABSTRACT	iii
	CONTENTS	iv
	FIGURES	v
	TABLES	vi
1.	INTRODUCTION	1
1.1	Reliability of Radioactivity Measurements	1
1.2	Background of the Hong Kong Observatory's participation in inter-laboratory comparison exercises	1
1.3	Objectives of the Study	2
2.	PERFORMANCE INDICATORS	3
2.1	Decile (D)	3
2.2	Ranking-score (RS)	3
2.3	Ratio (R)	4
2.4	z-score (Z)	5
2.5	Coefficient of Variation (CV)	5
2.6	Other Indicators of Performance	6
3.	RESULTS ABOUT HKO'S PERFORMANCE	7
3.1	Data	7
3.2	Decile	7
3.3	Ranking-score	7
3.4	Ratio	8
3.5	z-score	8
3.6	Coefficient of Variation	9
4.	CONCLUSIONS	11
	ACKNOWLEDGEMENTS	12
	REFERENCES	13

FIGURES

Page

Figure 1. Distribution of HKO's Results in the 'Decile' indicator 16

Figure 2. Ranking-score of HKO's Alpha Measurement Results 17

Figure 3. Ranking-score of HKO's Beta Measurement Results 18

Figure 4. Ranking-score of HKO's Gamma Measurement Results 19

Figure 5. Z-score of HKO's Alpha Measurement Results 20

Figure 6. Z-score of HKO's Beta Measurement Results 21

Figure 7. Z-score of HKO's Gamma Measurement Results 22

Figure 8. Coefficient of Variation of HKO's Alpha Measurement Results 23

Figure 9. Coefficient of Variation of HKO's Beta Measurement Results 24

Figure 10. Coefficient of Variation of HKO's Gamma Measurement Results 25

Figure 11. Homogeneity of IAEA-385 Samples (Cs-137 measurements) 26

TABLES

Page

Table 1. Summary of HKO's Participation in Inter-laboratory Comparison Exercises 27

Table 2. Summary of HKO's Performance by Ranking-score (RS) 28

Table 3. Summary of HKO's Performance by Ratio (R) 28

Table 4. Summary of HKO's Performance by Coefficient of Variation (CV) 29

1. INTRODUCTION

1.1 Reliability of Radioactivity Measurements

The reliability of radioactivity measurements depends on two factors, namely precision and accuracy. "Precision" is a measure of the repeatability of measurements, reflecting the closeness of agreement of independent test results (ISO 1997). It can be monitored by internal checking such as replicate measurements. "Accuracy" is a measure of the closeness of the result to the 'true' value (Thompson and Wood 1993). In environmental sample measurements, the 'true' values are often unknown and determination of accuracy of results relies on the implementation of quality assurance measures. Such measures include determination of the activity of the same sample by independent methods, analysts and techniques; control measurements of standard reference materials issued by international renowned institutes; and participation in inter-laboratory comparison exercises.

Inter-laboratory comparison exercises are widely used to determine the performance of radiation laboratories in conducting radioactivity measurements. These exercises provide a common platform enabling the performance of participating laboratories to be assessed objectively (IAEA 1989, ISO 1997).

For each inter-laboratory comparison exercise, randomly selected sub-samples from a source of material are distributed by the exercise organizer to the participating laboratories. These laboratories will then carry out measurements and report the results back to the organizer. Subsequently, the organizer dispatches an "assigned" value for the sample material obtained from statistical evaluations of the results from all the participating laboratories. The agreement of a laboratory's result with the assigned value thus constitutes a measure of the performance of the respective laboratory.

1.2 Background of the Observatory's participation in inter-laboratory comparison exercises

The Hong Kong Observatory (HKO) started radioactivity measurements in 1961. In response to the construction of nuclear power stations at Daya Bay in Guangdong, the HKO embarked in 1983 on a comprehensive Environmental Radiation Monitoring Programme (ERMP) to monitor the environmental radiation level in Hong Kong (HKO 1992). Environmental samples and foodstuff commonly consumed by the Hong Kong public are collected and analyzed at the King's Park Radiation Laboratory (KPRL) of the HKO. The details of the ERMP and the analysis results are given in the programme's series of annual reports (HKO 2005). HKO's participation in international inter-laboratory comparison exercises followed from the inception of ERMP.

Besides participating in international inter-laboratory comparison exercises, HKO also joined a number of regional inter-laboratory comparison exercises, split-sample analysis as well as proficiency tests (PT) [See Footnote below]*. These included the ones organized by the U.K. Defence Radiological Protection Service between 1988 and 1989 (Wong 1990), the Ministry of Agriculture, Fisheries and Food of the United Kingdom in 1991, the South China Sea Environmental Monitoring Centre in 1992, the Zhejiang Province Environmental Radiation Monitoring Centre and the Qinshan Nuclear Power Company in 1992 and 1995. HKO also participated in three PTs organized by the International Atomic Energy Agency (IAEA) between 1999 and 2002, as well as two comparison exercises on emergency radiation monitoring organized by the China Institute for Radiation Protection between 1999 and 2004. As the above exercises and tests involved a limited number of participating laboratories or employed different analysis schemes, the results were not included in this study.

1.3 Objectives of the Study

The study is based on the results of HKO in 17 inter-laboratory comparison exercises between 1989 and 2005. A summary of the exercises is given in Table 1. The objectives of the study are:

(i) to use objective indicators to evaluate the performance of the measurements at KPRL;
(ii) to evaluate KPRL's performance in the comparison exercises; and
(iii) to identify improvement areas.

Footnote* - "Proficiency tests" are referred to by IAEA as the exercises involving spiked samples with known activities to gauge the performance of participating laboratories (IAEA 2005b). They are called "known-value schemes" in ISO Guide 43-1 (1997) and are different from the inter-laboratory comparison exercises described in this report.

2. PERFORMANCE INDICATORS

A number of indicators are used in inter-laboratory comparison exercises like those organized by AQCS of IAEA (IAEA 2005a) and QUASIMEME of European Union (Cofino and Wells 1994) to gauge the performance of the participating laboratories. Some of these indicators evaluate the performance of a laboratory with respect to other participating laboratories while others assess the deviation of a laboratory's result relative to the assigned value.

To assess the performance of KPRL in the inter-laboratory comparison exercises, this report employs 5 different performance indicators. Similar versions of some of them are used by international exercise organizers like IAEA (IAEA 2005a; Thompson and Wood 1993). The indicators include "Decile", "Ranking-score", "Ratio", "z-score" and "Coefficient of variation". Each of them has its own characteristics and serves different purposes. Their definitions are described below.

2.1 Decile (D)

All results reported by the participating laboratories in an exercise are divided into 10 deciles in an ascending order of activity concentration. For example, a result will fall in the D^{th} Decile if it is the i^{th} result in ascending order of activity concentration such that:

$$D - 1 < 10 * i/n \leq D$$

where n is the total number of reported results.

Those results close to the median would then fall in the 5^{th} Decile or the 6^{th} Decile, while those results fell in the 1^{st} Decile or the 10^{th} Decile would likely be less than satisfactory.

This indicator shows how the result of a laboratory compares with those of other laboratories in a particular round of exercise. It is robust in the sense that the Decile of a result will not be affected by the result's actual value. It cannot show the performance of individual determination in an exercise precisely, but it provides a general understanding of the distribution of all the results over the years. Having results that are persistently in the upper deciles or the lower deciles indicates a possible systematic bias in the measurements.

2.2 Ranking-score (RS)

Rank (r) is a measure of a laboratory's performance based on the deviation of its result from the assigned value provided by the exercise organizer. Rank 1 is assigned to the result with the smallest deviation from the assigned value; Rank 2 is assigned to the result with the second smallest deviation; and so forth. That means a result having Rank 1 is the best while that having Rank n is the worst, where n is the total number of reported results.

"Ranking-score" (RS) normalizes the rank into a score. The RS for Rank 1 is assigned to be 100 and that for Rank n to be 0, with those in between being equally divided among themselves. The RS of a result is defined as follows:

$$RS = 100 \times (n-r) / (n-1)$$

where r is the Rank of the reported result; and
n is the number of reported results.

RS is not an indication of the closeness to the assigned value but rather the percentage of laboratories with performance better or worse than the laboratory in an exercise.

Similar to Decile, the RS over time of a laboratory may not serve as a proper indicator for monitoring the performance trend of a laboratory as different rounds of exercises may have different participants. The indicator should also be interpreted with caution (ISO 1997), particularly when the number of participating laboratories is too few. For example, a laboratory second in ranking has RS of only 50 in an exercise involving three participants, but will have RS of 75 in an exercise involving five participants.

2.3 Ratio (R)

The Ratio R is defined as the ratio of the reported value to the assigned value. This indicator is similar to the Q-score described by Thompson and Wood (1993).

$$R = V_{reported} / t$$

where $V_{reported}$ is the reported value; and
t is the assigned value provided by the exercise organizer.

It is a simple measure of agreement between the measured and assigned activity concentrations without reference to the performance of other laboratories. The closer the value of R approaches unity, the better will be the agreement between the result reported by a laboratory and the assigned value. For illustration purpose, results with R within 10 percent from unity are categorized as very good while those within 20 percent from unity as satisfactory (Moore et al. 1989). Results outside 50 percent will require investigation.

This indicator reflects the difference between a laboratory's result and the assigned value. However, in some cases the value of R alone does not truly reflect the performance of the laboratory. For instance, R is not symmetrical with respect to unity, i.e. it can only differ from unity by 1 at the low end, but any value at the high end. Furthermore, samples with low activity concentrations usually have relatively large uncertainties (Horwitz 1982) and may lead to large values of R for many participating laboratories. Thus other indicators have to take into account when interpreting results with R deviating significantly from unity.

2.4 z-score (Z)

The z-score Z is defined by the following equation:

$$Z = (V_{reported} - t)/SD$$

where $V_{reported}$ is the reported value;
t is the assigned value provided by the exercise organizer; and
$SD = [\ \Sigma(V_i - t)^2 / n\ ]^{1/2}$ is the standard deviation (with respect to t) of all reported values, V_i, with outliers designated in the exercise reports excluded.

z-score is commonly used in evaluating the performance of a laboratory (Cofino and Wells 1994, ISO 1997; IAEA 2005a). It is a simple method giving a normalized performance score based on the difference between the laboratory's result and the assigned value. The z-score for a result indicates how far and in what direction that result deviates from the assigned value. In a comparison exercise the performance of a laboratory is usually considered to be satisfactory when $|Z| \leq 2$ (ISO 1997), which represents the range that 95% of reported results would fall within.

Since the SD is calculated based on results from other laboratories in the exercise, the z-score also provides an indication of the performance of a laboratory with respect to other laboratories in the same exercise, in additional to the difference between the reported value and the assigned value. However, since the participating laboratories may not be the same in different rounds of exercises, comparing z-scores obtained from different exercises requires extra caution (Thompson and Wood 1993).

In some exercises, a predetermined value, from past experience, estimates, or other means (World Health Organisation 1996), is assigned to the SD. In this case, the "absolute" performance of a laboratory can be gauged and the z-score so obtained can be plotted in the form of a time series to show the variation of performance of a laboratory over time (Thompson and Wood 1993).

2.5 Coefficient of Variation (CV)

The coefficient of variation CV is defined as the standard deviation SD (with respect to the assigned value t) of the measurements by a laboratory divided by the assigned value t, expressed in percentage (Moore et al. 1989), i.e.:

$$CV = SD/t \times 100\ \%$$

where $SD = [\ \Sigma(x_i - t)^2 / m\ ]^{1/2}$;
x_i is the i^{th} measured value; and

m is the number of measurements carried out by that laboratory.

The coefficient of variation is a measure of the spread of results from repeated measurements by a laboratory. It is useful in tracking the dispersion in measurement results of a laboratory and identifying possible mistakes and errors during the measurement process.

2.6 Other Indicators of Performance

Other indicators of performance are also used in some proficiency tests, such as the "u-test" score (IAEA 2005b) and E_n numbers (ISO 1997; Standards Council of Canada 2001). The u-test score and E_n number are similar[See Footnote below]*, defined as the ratio of the difference between the reported result and the assigned value to the combined uncertainty. For instance, the u-test score is defined as follows:

$$U = |(V_{reported} - t)| / (\sigma_{reported}^2 + \sigma_t^2)^{½}$$

where U is the u-test score;

$V_{reported}$ is the reported value;

t is the assigned value provided by the exercise organizer;

$\sigma_{reported}$ is the combined standard uncertainty of the reported value; and

σ_t is the combined standard uncertainty of the assigned value provided by the exercise organizer.

This score incorporates in its calculation the uncertainties associated with the reported result and the assigned value. It is usually used as an acceptance criterion in proficiency tests. As the uncertainty for the assigned value is not available in reports of inter-laboratory comparison exercises, those two indicators, i.e. the u-test score and E_n number, were not included in this study.

*Footnote: The formula of E_n is similar to that for the u-test score except E_n carries a sign.

3. RESULTS ABOUT HKO'S PERFORMANCE

3.1 Data

HKO's results in terms of the 'Decile' indicator are plotted in Figure 1. The ranking-score of HKO's results regarding alpha, beta and gamma radiation measurements are shown in Figures 2, 3 and 4 respectively. The z-scores are shown in Figures 5, 6 and 7 while the coefficients of variation are presented in Figures 8, 9 and 10. Summaries of HKO's performance are tabulated in Tables 2 to 4.

In presenting the results, each reported radionuclide was counted as a separate result, referred to as a "determination". For example, in a gamma measurement, if both Cs-137 and K-40 were reported, it would be counted as two determinations of gamma activity.

3.2 Decile

It was observed from Figure 1 that the alpha results were evenly distributed between the lower and the upper deciles - 50% of the determinations in the lower deciles and 50% of the determination in the upper deciles. No results fell in Decile 1 nor Decile 10.

For the beta results, 67% of the determinations were in the lower deciles with three determinations in Decile 1. Ignoring those three determinations, other results were evenly distributed around the middle. Those three less-than-satisfactory determinations in respect of this indicator were the determinations of Sr-90 in clover IAEA-156 (IAEA 1991a), in soil IAEA-327 (IAEA 2001) and in sea water IAEA-381 (IAEA 1999). Even though the results for IAEA-156 and IAEA-327 were treated as less-than-satisfactory for the 'Decile' indicator in this report, the magnitude of their z-scores were less than 2 and were considered acceptable by the organizer. The case for IAEA-381 will be discussed in more details in Section 3.5.

In respect of the gamma measurements, 63% of the determinations fell in the upper deciles while the remaining 37% fell in the lower deciles. There was no determination in Decile 1 but one determination in Decile 10. This less-than-satisfactory result for this indicator was the determination of Cs-137 in sea sediment IAEA-385 (IAEA 2005a). However, this result was accepted by the organizer as the magnitude of the z-score for this determination was less than 2. More detailed discussion on this measurement will be made in Section 3.5.

3.3 Ranking-score

A summary of the performance of HKO by ranking-score is given in Table 2, while the distributions of the ranking-score for the alpha, beta and gamma measurements are shown in Figures 2 to 4 respectively. The average ranking-scores for alpha, beta and gamma measurements were 61, 51 and 57 respectively. HKO's performance was in general satisfactory for all the three types of measurements. The overall ranking-score of 57 also put HKO among the better performing group of the participating laboratories.

3.4 Ratio

From the overall result given in Table 3, nearly 60 % of all results were very good, i.e. within 10 percent of the assigned value; and around 90 % were satisfactory (within 20 percent of the assigned value). The performance of the alpha measurements was the best among the three types of measurements with all results being satisfactory or better. More than half (56%) of the beta results were rated satisfactory or better. For gamma measurements, nearly 90% of the results were satisfactory or better.

There were two beta results with R outside 50% of the assigned value. They were the determinations of Sr-90 in sea water IAEA-381 (IAEA 1999) and in lagoon sediment IAEA-384 (IAEA 2000). As remarked in Section 2.3, this indicator has to be interpreted with caution since the value of R does not take account of other important factors such as measurement uncertainties. The case for IAEA-384 was that while R was equal to 2.13 (the largest among the R values for beta measurement) yet the z-score of the reported result was less than 2, well within IAEA's range of acceptable results. The case for IAEA-381 will be discussed in more details in Section 3.5.

3.5 z-score

The z-scores for the alpha, beta and gamma measurements are shown in Figures 5, 6 and 7 respectively. In general, the performance of a determination was considered satisfactory if $|Z| <= 2$ (ISO 1997). There were only one beta determinations and two gamma determinations that were less-than-satisfactory, i.e. with the magnitude of z-score greater than 2.

The less-than-satisfactory case for beta measurement was the determination of Sr-90 in sea water IAEA-381 (IAEA 1999) with Z=-4.41. The case was likely due to the high contents of naturally occurring non-radioactive metal ions such as Ca^{2+}, Na^{+} and Sr^{2+} in the sample, which were found to have interfered with the chemical treatment for Sr-90. This had an effect on the gravimetric determination of the chemical yield and led to an inaccurate determination of the activity concentration.

The two less-than-satisfactory cases for gamma measurement were the determination of K-40 in marine cockle flesh IAEA-134 (IAEA 1993a) with Z=3.51 and the determination of Cs-137 in sea fish IAEA-414 (IAEA 2004) with Z=2.14. In respect of the former, K-40 is a naturally occurring primordial radionuclide existing practically everywhere and provides a good benchmark in practically all gamma measurements of environmental samples. On the basis of this, the use of a non-updated file in background subtraction was identified to be the major cause for the less-than-satisfactory result.

In the case of Cs-137 determination in IAEA-414 the reason was less obvious, but it was considered that sample inhomogeneity might have played a vital part. The exercise organizer normally carried out homogeneity tests on the comparison sample before dividing it into a number of sub-samples for distribution to different laboratories (ISO 2006). Nonetheless, these sub-samples could still have natural variations (ICRU 2006). It was thus possible that the 'true' activity of a sub-sample received at a laboratory might be quite different from the assigned value, even though the assigned value was representative of a large batch of sub-samples.

In the IAEA-385 exercise in 2002 (IAEA 2005a), the z-scores of HKO's reported values for K-40 and Cs-137 were close to 2. Investigations were carried out to identify possible improvements for the exercise. On the advice of IAEA, it was recognized that there might be variations in the activities of the sub-samples for different laboratories. Figure 11 shows that although HKO's reported value was relatively higher, yet it was very close to the range of those sub-samples in IAEA's homogeneity test. For further confirmation, a repeat measurement by a separate independent laboratory, the Government Laboratory (GL), was conducted. The z-score of the GL's result was 2.66, even higher than that of the HKO's. To conclude, it was highly likely that the difference between HKO's result and the assigned value might arise from possible inhomogeneity among the sub-samples distributed to the participating laboratories.

3.6 Coefficient of variation

From Table 4, it can be seen that 56% of all HKO results had CV less than 10%, and 81 % of HKO's results had CV less than 20%. The dispersion of HKO measurement results was in general acceptable, with the exception of only two determinations having CV greater than 50%. These two less-than-satisfactory determinations were U-235 measurement in marine cockle flesh IAEA-134 (IAEA 1993a), and Sr-90 measurement in sea water sample IAEA-381 (IAEA 1999) which was already discussed in Section 3.5 above.

The case of U-235 in IAEA-134 was found to be associated with the close interfering peaks of U-235 and Ra-226 around the energy of 186 keV. This led to the least satisfactory performance in CV hitherto (900%, the largest deviation in percentage from the assigned value) in all comparison exercises. This also compromised the overall performance over the years. It was noted that spectral analysis software was not always capable of resolving close interfering peaks and the use of direct software output would lead to erroneous results.

The CV for the three types of measurements over time are also plotted in Figures 8, 9 and 10. Though the linear fits might have slight increasing or decreasing slopes, t-tests indicated that there were no significant trends at the 95% confidence level for all three types of measurements. The data point for IAEA-134 U-235 measurement was not included in the trend analysis for gamma measurements as it would generate an overwhelming bias and an unrealistic decreasing trend for the whole data set.

4. CONCLUSIONS

The results of HKO's participation in inter-laboratory comparison exercises between 1989 and 2005 were reviewed. A number of objective indicators were used for assessing HKO's performance in the exercises.

The performance of HKO in inter-laboratory comparison exercises on average ranked among the group of better performing laboratories in the exercises. As a whole, there was no major fault or inadequacy in the equipment. Furthermore, no significant trend in the performance was identified at the 95% confidence level in all three types of measurements.

There were a handful of determinations which deserved attention on the basis of the various indicators. Of these determinations, only three have less-than-satisfactory z-scores, i.e. with magnitude greater than 2. Studies were carried out to find out the possible causes. The causes identified included: interference from high contents of non-radioactive isotopes in gravimetric determination of the chemical yield of a sample, errors introduced in the background subtraction during the activity determination of a naturally occurring primordial radionuclide, and possible inhomogeneity among sub-samples distributed to the participating laboratories.

The results shown in the figures and tables in this report give good indications of HKO's long term performance. They also help reveal possible biases not readily observable in a single exercise. Besides performance monitoring, another benefit of inter-laboratory comparison exercises was to provide opportunities to constantly review the laboratory procedures. The findings obtained in this study would serve as useful guidance and reference for future inter-laboratory comparison exercises as well as routine measurements in the ERMP for the monitoring of environmental and food samples.

ACKNOWLEDGEMENTS

We wish to thank Dr. M.K. Pham, Marine Environment Laboratory of International Atomic Energy Agency, Dr. B.Y. Lee, Dr. K.K. Yeung and Dr. C.M. Cheng of the Hong Kong Observatory for their valuable advice on the exercise results and the draft of the report. Special thanks also go to the International Atomic Energy Agency, World Health Organization and Zhejiang Province Environmental Radiation Monitoring Centre for organizing the inter-laboratory comparison exercises. We would also like to express our gratitude to the Government Laboratory for the assistance rendered in chemical treatment of samples.

REFERENCES

1.	Cofino, W.P. and D.E. Wells	1994	Design and evaluation of the QUASIMEME inter-laboratory performance studies: A test case for robust statistics, Marine Pollution Bulletin, V.29, p. 149-158.
2.	HKO	1992	Background Radiation Monitoring Programme 1987 – 1991, Technical Reports on Environmental Radiation Monitoring in Hong Kong No.8, Hong Kong Observatory (HKO).
3.	HKO	2005	Summary of Environmental Radiation Monitoring in Hong Kong 2004, Technical Reports on Environmental Radiation Monitoring in Hong Kong No.25, Hong Kong Observatory (HKO).
4.	Horwitz, W.	1982	Evaluation of Analytical Methods Used for Regulation of Foods and Drugs, Analytical Chemistry, Vol. 54, No. 1, p. 67A – 76A.
5.	ICRU	2006	ICRU Report 75 - Sampling for Radionuclides in the Environment, International Commission on Radiation Units and Measurements (ICRU).
6.	International Atomic Energy Agency	1989	Technical Reports Series No. 295: Measurement of Radionuclides in Food and the Environment – A Guidebook.
7.	International Atomic Energy Agency	1990	Intercomparison of Radionuclide Measurements in Mediterranean Tuna Fish Sample IAEA-352.
8.	International Atomic Energy Agency	1991a	Intercomparison Study IAEA-156 on the Determination of Radionuclides in Clover.
9.	International Atomic Energy Agency	1991b	Intercomparison of Radionuclide Measurements in Marine Sediment Sample IAEA-367.
10.	International Atomic Energy Agency	1993a	Intercomparison of Radionuclide Measurements in Marine Cockle Flesh Sample IAEA-134.

11.	International Atomic Energy Agency	1993b	Intercomparison of Radionuclide Measurements in Marine Sediment Sample IAEA-135.
12.	International Atomic Energy Agency	1994	Intercomparison of Radionuclide Measurements on Marine Sediment Sample IAEA-300.
13.	International Atomic Energy Agency	1996	Intercomparison Run IAEA-375 : Determination of Radionuclides in Soil Sample IAEA-375.
14.	International Atomic Energy Agency	1997	Intercomparison of Radionuclide Measurements on Marine Sediment IAEA-315.
15.	International Atomic Energy Agency	1999	Intercomparison of Radionuclide Measurements in Irish Sea Water IAEA-381.
16.	International Atomic Energy Agency	2000	Intercomparison Run IAEA-384 : Radionuclides in Fangatafa Lagoon Sediment.
17.	International Atomic Energy Agency	2001	Intercomparison Run IAEA-326 and IAEA-327 : Determination of Radionuclides in Soils.
18.	International Atomic Energy Agency	2004	Intercomparison IAEA-414 : Radionuclides in Mixed Fish from Irish Sea and the North Sea.
19.	International Atomic Energy Agency	2005a	Intercomparison Exercise IAEA-385 : Radionuclides in Irish Sea Sediment.
20.	International Atomic Energy Agency	2005b	Final Report on the Proficiency Test of the Analytical Laboratories for the Measurement of Environmental Radioactivity (ALMERA) Network, IAEA/AL/152.
21.	ISO	1997	Proficiency testing by interlaboratory comparisons – Part 1: Development and operation of proficiency testing schemes, Guide 43-1.
22.	ISO	2006	Reference Materials – General and Statistical principles for certification, Guide 35.

23.	Moore, J., J. Broadway and R. Blandchard	1989	The Eastern Environmental Radiation Facility Participation in Interlaboratory and Intralaboratory Comparisons of Environmental Sample Analyses : 1981 – 1986, Office of Radiation Programs, U. S. Environmental Agency.
24.	Standards Council of Canada	2001	Proficiency testing by interlaboratory comparison – Part 1: Development and operation of proficiency testing schemes
25.	Thompson, M. and R. Wood	1993	The international harmonized protocol for the proficiency testing of (chemical) analytical laboratories, Pure & Appl. Chem. v. 65, no. 9, p. 2123-2144.
26.	Wong, M.C.	1990	An Inter-Laboratory Comparison between Measurements of Radioactivity in Seabed Sediments by the Hong Kong Observatory Hong Kong and the United Kingdom Defence Radiological Protection Service, Hong Kong Observatory, Technical Reports on Environmental Radiation Monitoring in Hong Kong No.2.
27.	World Health Organisation	1996	Report to Participants in GEMS/Food - EURO Proficiency Testing Exercise 95/01, Central Science Laboratory.

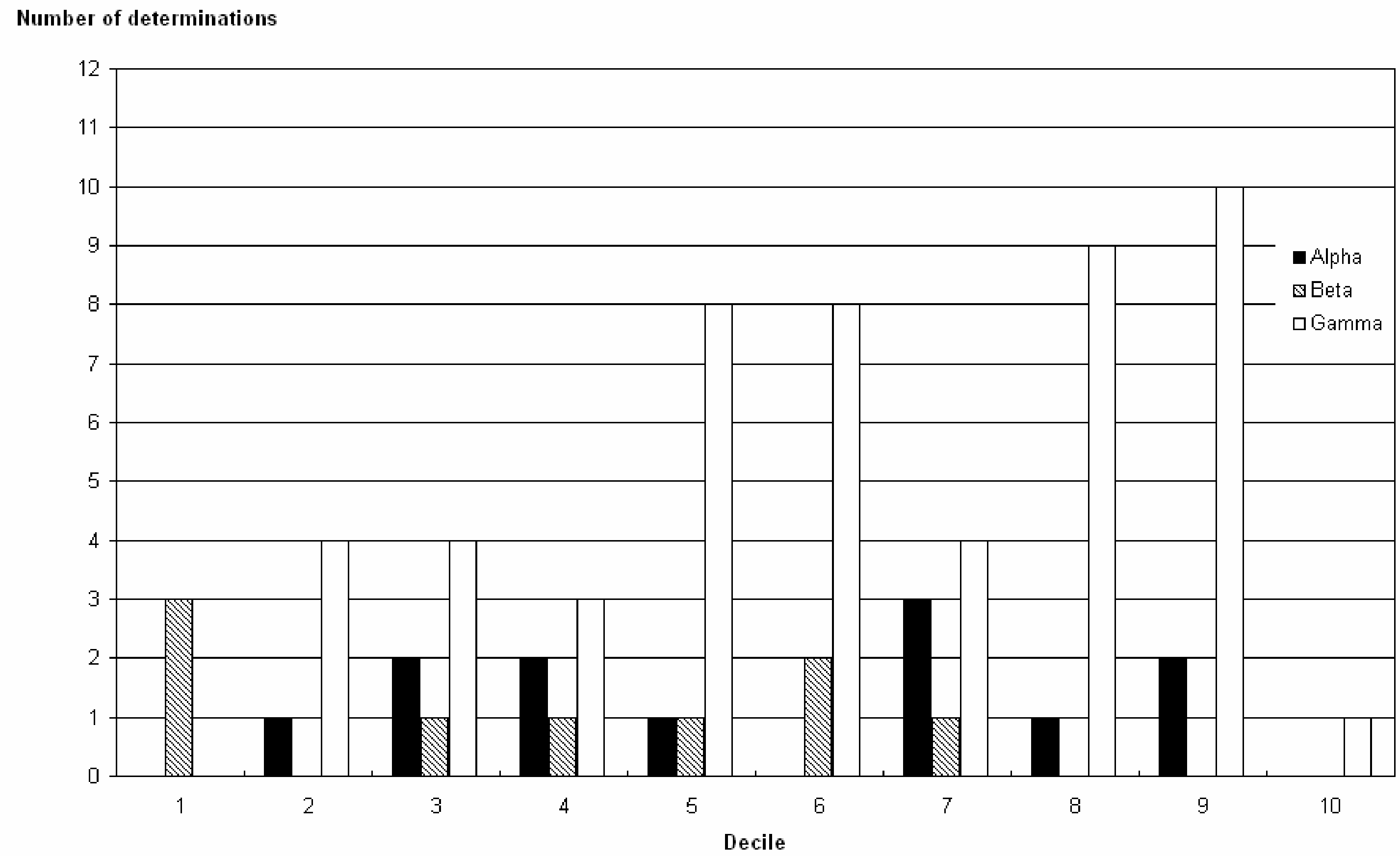


Figure 1. Distribution of HKO's Results in the 'Decile' indicator

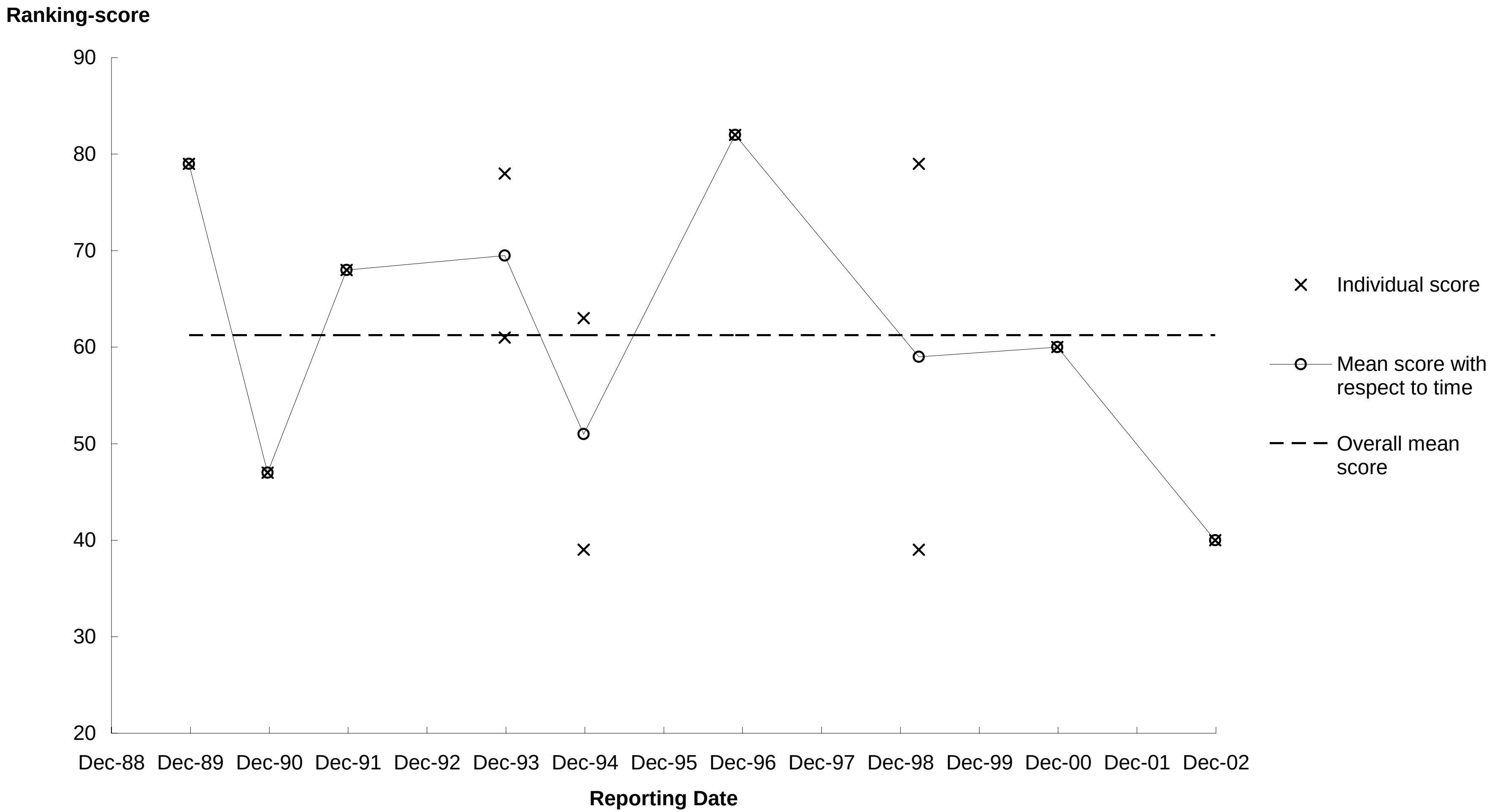


Figure 2. Ranking-score of HKO's Alpha Measurement Results

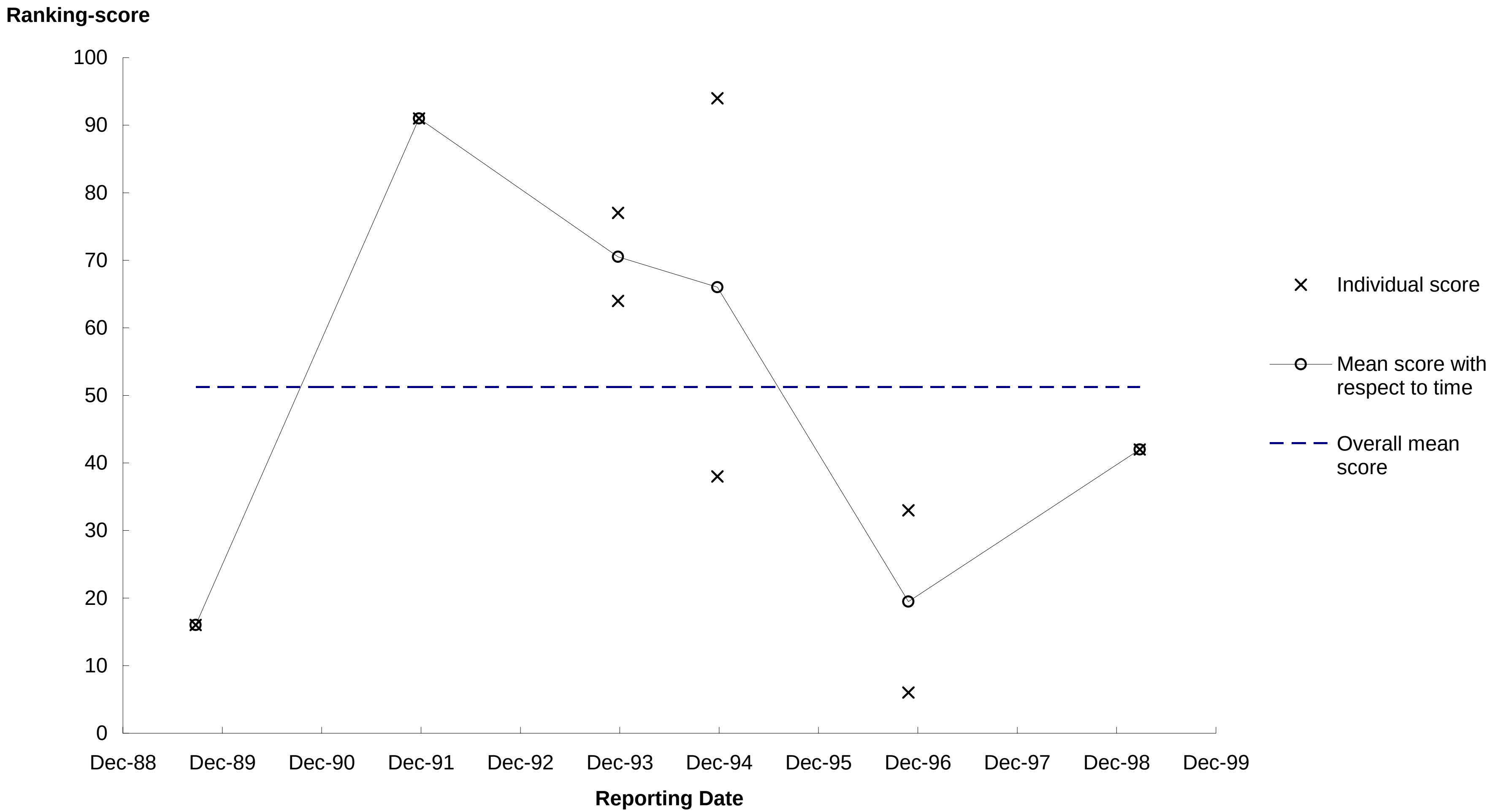


Figure 3. Ranking-score of HKO's Beta Measurement Results

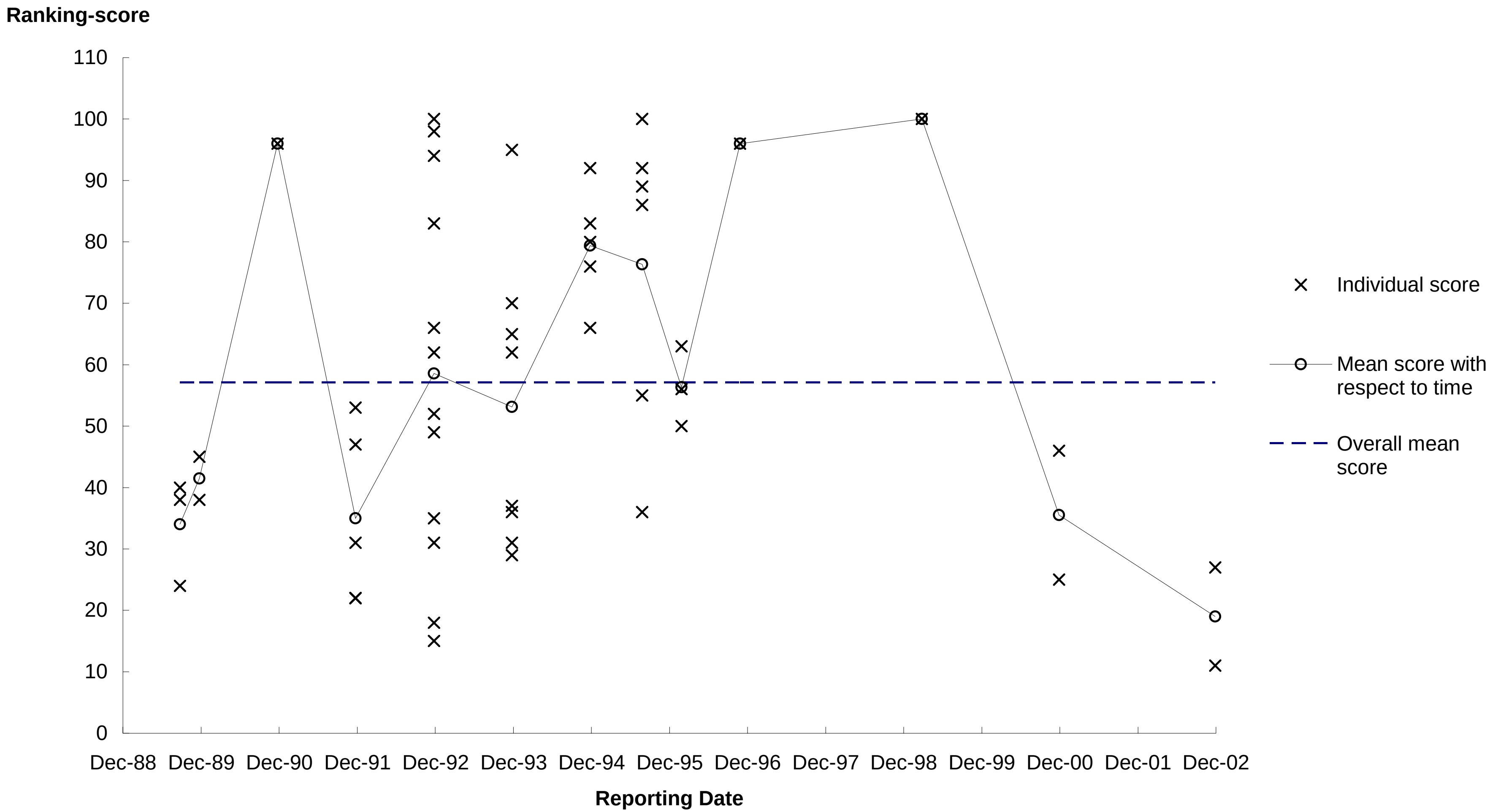


Figure 4. Ranking-score of HKO's Gamma Measurement Results

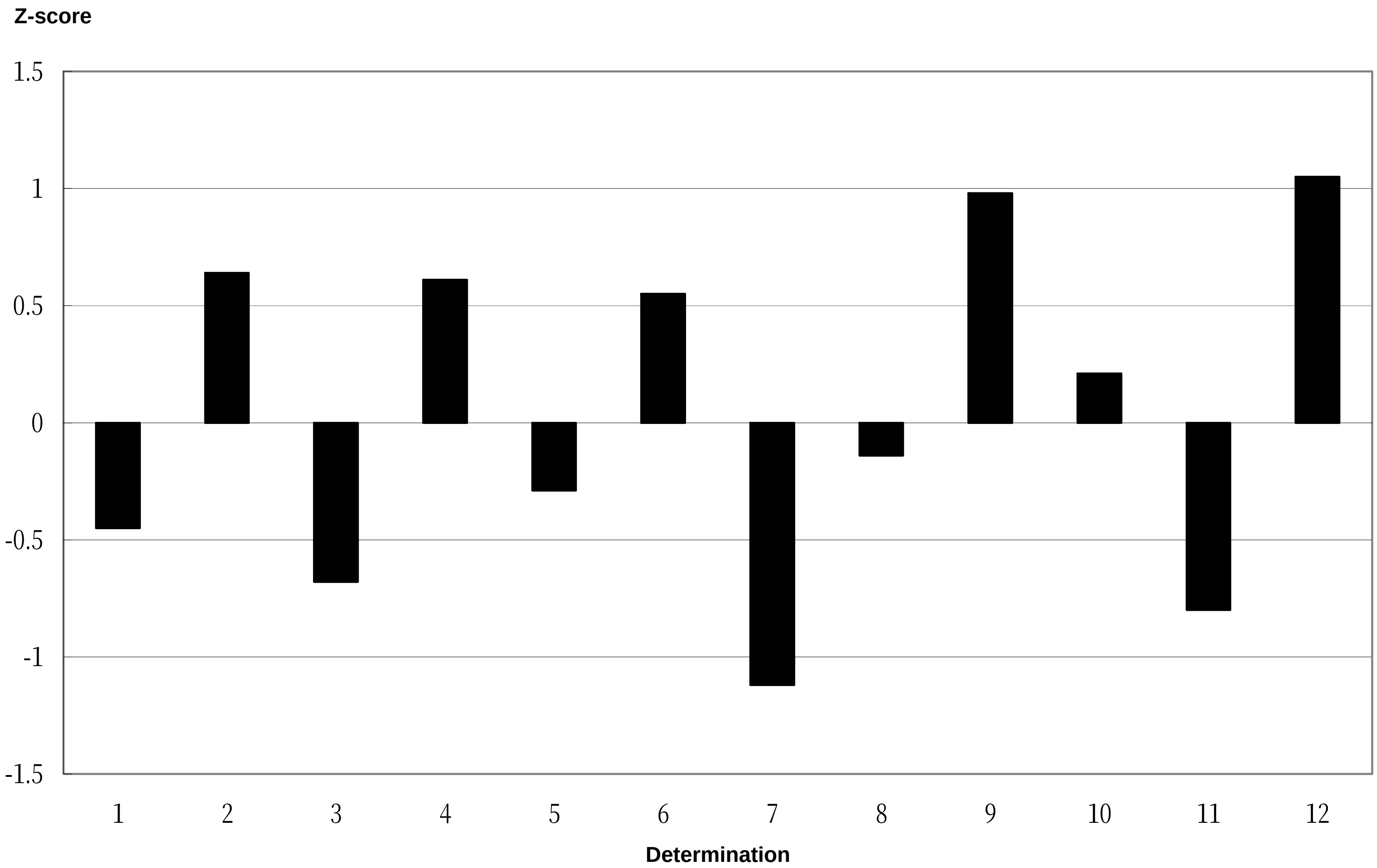


Figure 5. *Z*-score of HKO's Alpha Measurement Results

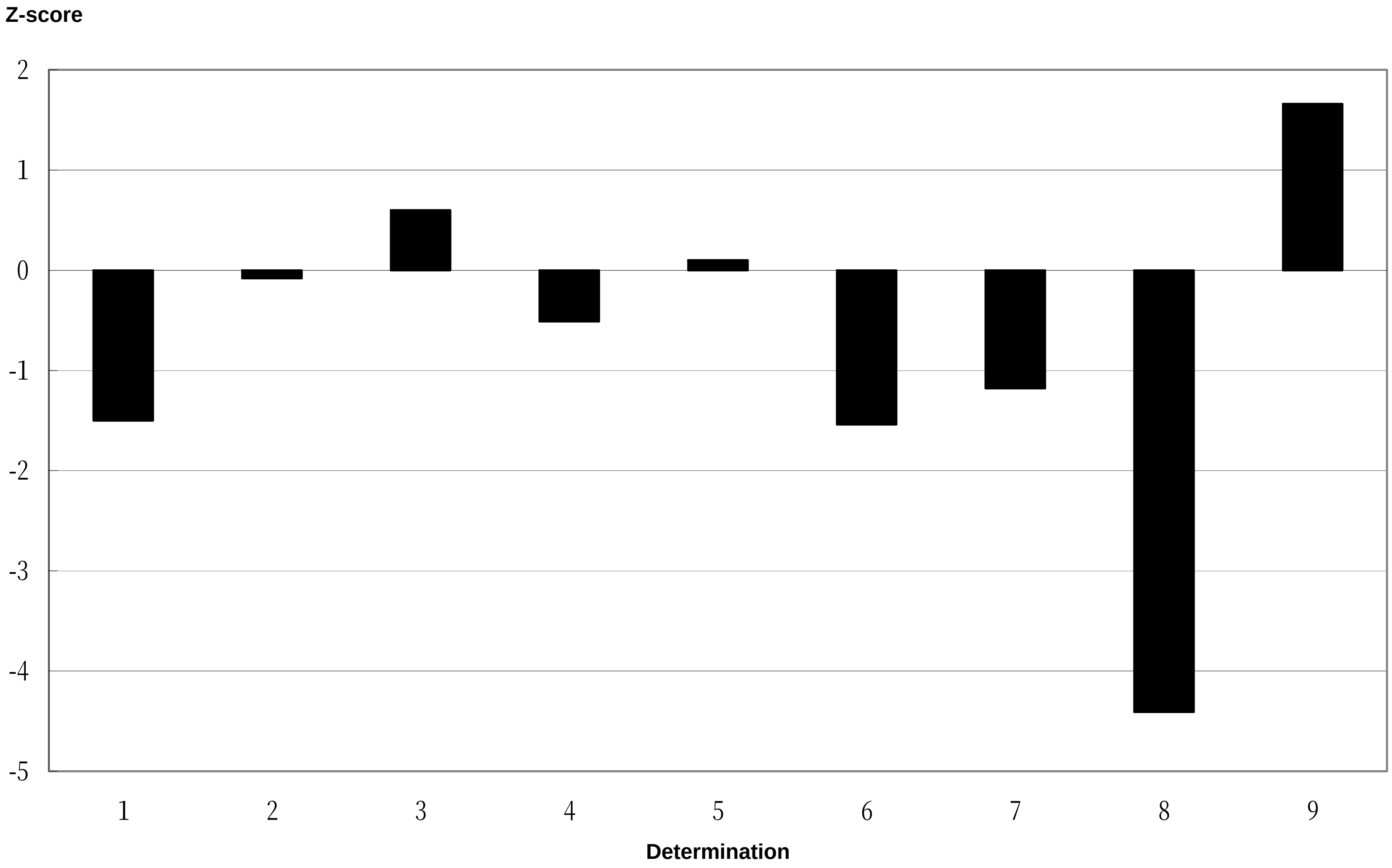


Figure 6. Z-score of HKO’s Beta Measurement Results

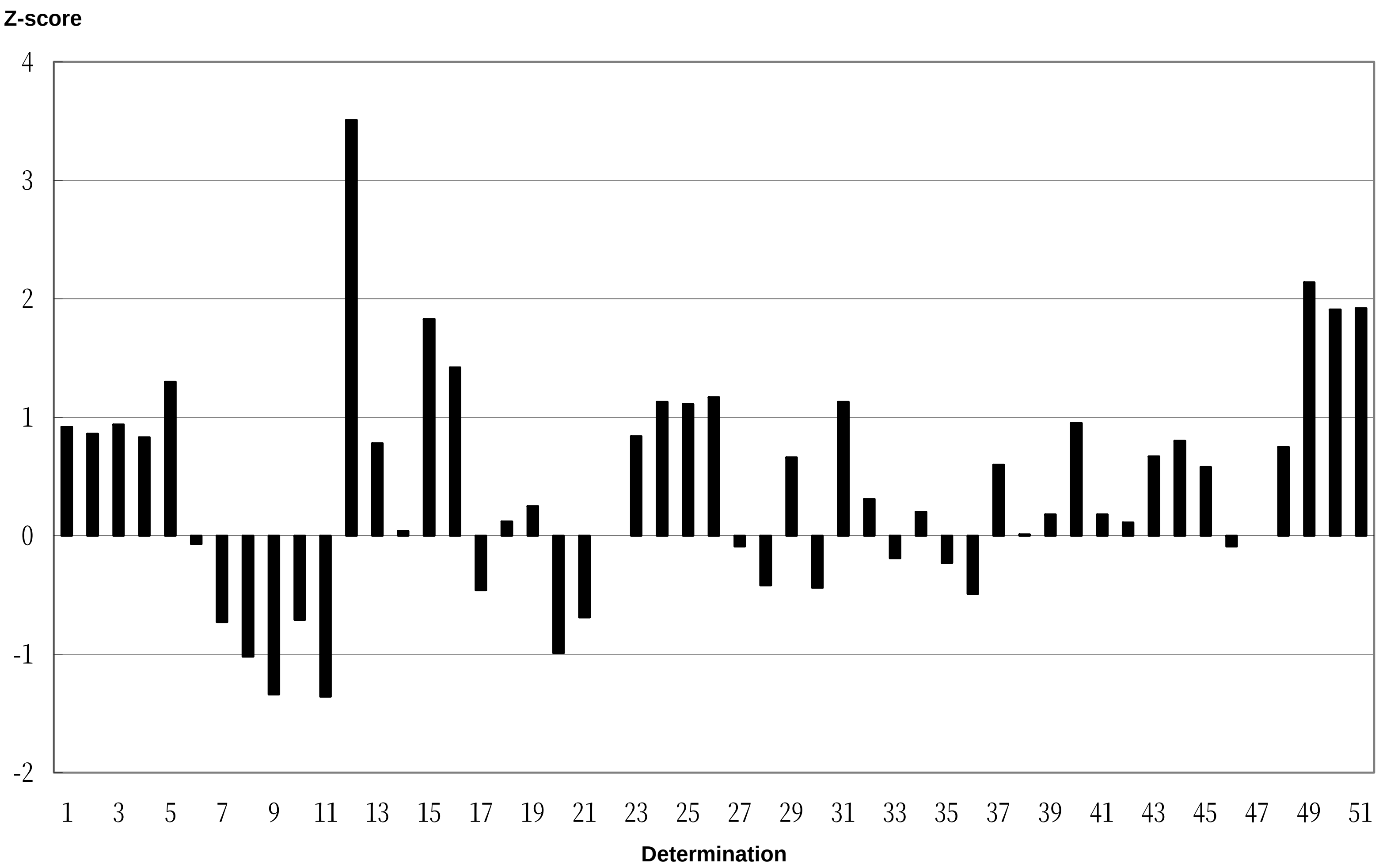


Figure 7. Z-score of HKO's Gamma Measurement Results

Figure 8. Coefficient of Variation of HKO's Alpha Measurement Results

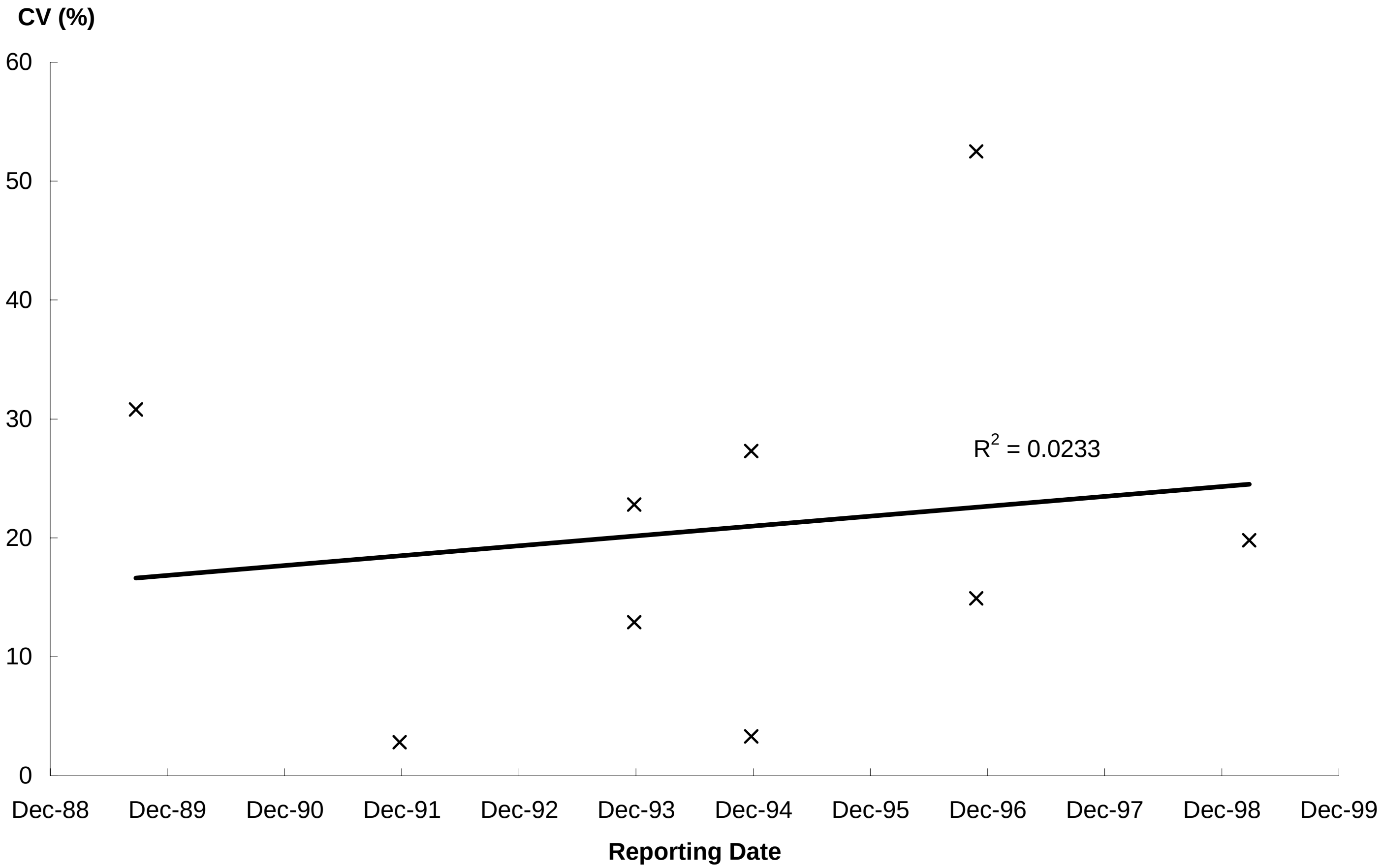
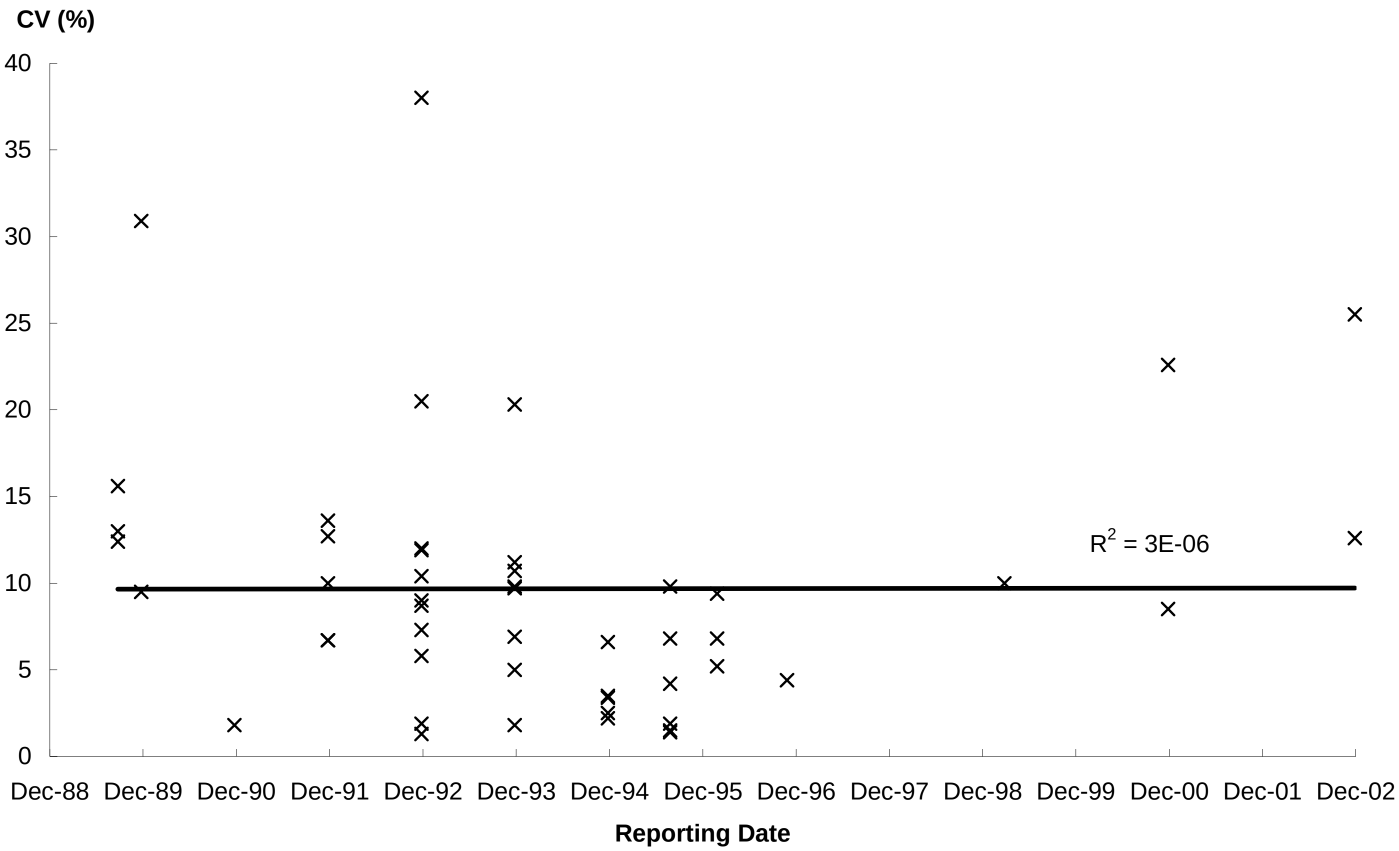
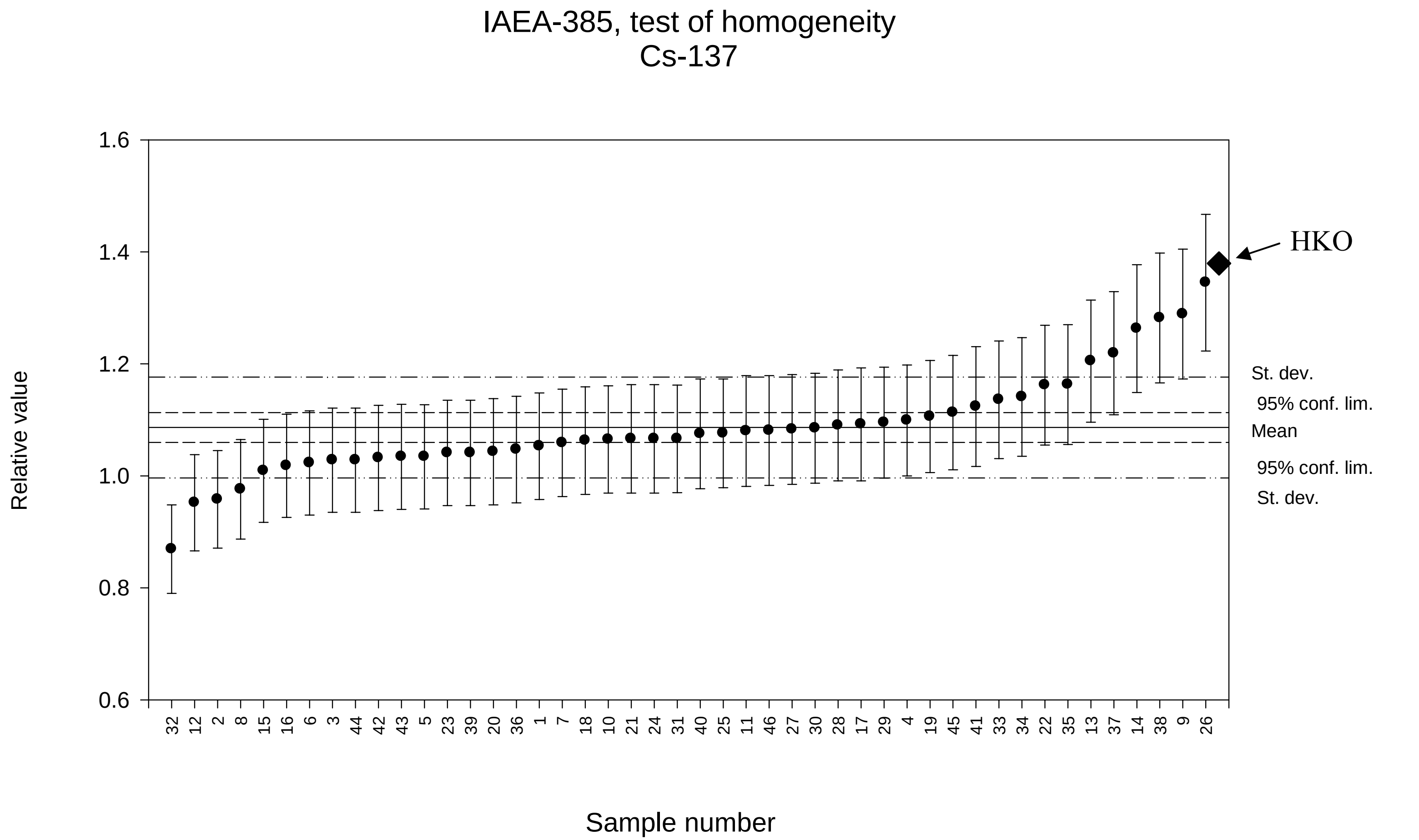


Figure 9. Coefficient of Variation of HKO’s Beta Measurement Results



Remark: The determination of U-235 in IAEA-134 (CV=900%) is not included in the Figure.

Figure 10. Coefficient of Variation of HKO's Gamma Measurement Results



Note: The above figure is extracted from Report on the worldwide Intercomparison Exercise IAEA-385 (IAEA 2005a).

Figure 11. Homogeneity of IAEA-385 Samples (Cs-137 measurements)

Table 1. Summary of HKO's Participation in Inter-laboratory Comparison Exercises

Reporting date	Organizing agency *	Comparison sample	Types of measurement performed	Results publishing date
September 1989	IAEA	IAEA-156 Clover	beta (Sr-90), gamma	January 1991 (IAEA 1991a)
December 1989	IAEA	IAEA-352 Tuna Fish	alpha (Po-210), gamma	August 1990 (IAEA 1990)
December 1990	IAEA	IAEA-367 Pacific Ocean Sediment	alpha (Pu-239/240), gamma	August 1991 (IAEA 1991b)
December 1991	IAEA	IAEA-375 Soil	alpha (Pu-239/240), beta (Sr-90), gamma	February 1996 (IAEA 1996)
December 1992	IAEA	IAEA-134 Marine Cockle Flesh	gamma	August 1993 (IAEA 1993a)
December 1992	IAEA	IAEA-135 Irish Sea Sediment	gamma	August 1993 (IAEA 1993b)
December 1993	IAEA	IAEA-300 Baltic Sea Sediment	alpha (Pu-239/240), beta (Sr-90), gamma	September 1994 (IAEA 1994)
December 1993	IAEA	IAEA-315 Arabian Sea Sediment	alpha (Pu-239/240), beta (Sr-90), gamma	December 1997 (IAEA 1997)
December 1994	IAEA	IAEA-326 Black Soil	alpha (Pu-239/240), beta (Sr-90), gamma	April 2001 (IAEA 2001)
December 1994	IAEA	IAEA-327 Podsolic Soil	alpha (Pu-239/240), beta (Sr-90), gamma	April 2001 (IAEA 2001)
August 1995	WHO	95/01 Milk Powder A	gamma	April 1996 (WHO 1996)
August 1995	WHO	95/01 Milk Powder B	gamma	April 1996 (WHO 1996)
February 1996	ZPERMC	141 Cylindrical Sediment Sample	gamma	March 1997
November 1996	IAEA	IAEA-381 Irish Sea Water	alpha (Pu-239/240), beta (H-3, Sr-90), gamma	1999 (IAEA 1999)
March 1999	IAEA	IAEA-384 Fangataufa Lagoon Sediment	alpha (Pu-238, Pu-239/240), beta (Sr-90), gamma	2000 (IAEA 2000)
December 2000	IAEA	IAEA-414 Irish and North Sea Fish	alpha (Pu-239/240), beta (Sr-90), gamma	2004 (IAEA 2004)
December 2002	IAEA	IAEA-385 Irish Sea Sediment	alpha (Pu-239/240), gamma	2005 (IAEA 2005a)

Note:

* -- IAEA - International Atomic Energy Agency
WHO - World Health Organization
ZPERMC - Zhejiang Province Environmental Radiation Monitoring Centre

Table 2. Summary of HKO's Performance by Ranking-score (RS)

Type of measurement	Number of determinations	Average score	Standard deviation
alpha	12	61	17
beta	9	51	32
gamma	51	57	27
overall	***72***	***57***	***26***

Table 3. Summary of HKO's Performance by Ratio (R)

Type of measurements	Number of determinations	Range in R	Mean value and standard deviation of R	Number of values differing from R = 1 by :		
				<= 10 %	<= 20 %	> 20 %
alpha	12	0.81 - 1.11	0.99 (0.10)	7 (58%)	12 (100%)	0 (0%)
beta	9	0.48 – 2.13	1.00 (0.47)	2 (22%)	5 (56%)	4 (44%)
gamma	51	0.88 - 10	1.22 (1.16)	35 (69%)	46 (90%)	5 (10%)
overall	***72***	***0.48 - 10***	***1.16 (1.07)***	***44 (61%)***	***63 (88%)***	***9 (12%)***

Table 4. Summary of HKO's Performance by Coefficient of Variation (CV)

Type of measurements	Number of determinations	Range in CV (%)	Mean value and standard deviation of CV (%)	Number of values differing from CV by :		
				10 %	20 %	> 50 %
alpha	12	5.4 – 26.9	13.6 (7.6)	5 (42%)	9 (75%)	0 (0%)
beta	9	2.8 – 52.5	20.8 (15.3)	2 (22%)	5 (56%)	1 (11%)
gamma	51	1.3 - 900	27.1 (125)	33 (65%)	44 (86%)	1 (2%)
overall	***72***	***1.3 - 900***	***24.1 (105)***	***40 (56%)***	***58 (81%)***	***2 (3%)***