



Digitalization sponsored  
by Thünen-Institut

*ICES SIXTH ROUND INTERCALIBRATION  
FOR  
TRACE METALS IN ESTUARINE WATER  
JMG 6/TM/SW*

by

S.S. Berman and V.J. Boyko

Marine Analytical Chemistry Standards Program  
Division of Chemistry  
National Research Council  
Ottawa, Ontario, Canada K1A 0R9

This is an extract of the full report which has been  
submitted to ICES in the Cooperative Research Report  
series.

## INTRODUCTION

A request was made to the ICES Marine Chemistry Working Group (MCWG) by the Joint Monitoring Group (JMG) of the Oslo and Paris Commissions in 1985 for the conduct of an intercalibration exercise for trace metals in estuarine waters. This request followed an evaluation of the results of the ICES Fifth Round Intercalibration for Trace Metals in Seawater, ICES 5/TM/SW (Berman et. al., 1986). It was suggested that the low concentration levels of the metals in the coastal water used in that study might not be typical of the levels prevailing at most of the JMP coastal and estuarine monitoring stations.

The MCWG, at its annual meeting in February 1985 (MCWG, 1985), formulated a proposal for the conduct of an intercalibration exercise based on the use of low salinity water (10 to 20 parts per thousand) collected from a North Sea estuary. The National Research Council of Canada (NRCC) agreed to undertake this project under contract to the Oslo and Paris Commissions. This laboratory was chosen because of its extensive experience in collecting and filtering large quantities of seawater for the preparation of seawater reference materials and could carry out the proposed work within the shortest possible timescale and at lowest cost to the Commissions.

The proposal was approved by the Oslo and Paris Commissions in June 1985 (OSPAR, 1985) who noted that all laboratories submitting seawater data in the Joint Monitoring Program (JMP) were expected to participate in the exercise, that the exact location of the sampling station be chosen by the JMG and that the report be submitted directly to the JMG at its meeting in January 1987.

The detailed planning for the exercise on behalf of the JMG was done by Dr. Arne Jensen (then chairman of the JMG) from the Marine Pollution Laboratory in Denmark, aided by Dr. Wim Cofino of the Rijkswaterstaat, the Netherlands, and Dr. Shier Berman of NRCC who assumed the position of overall coordinator. Dr. Cofino also arranged for the use of a Netherlands ship, the M.V. Argus, and onshore laboratory facilities in Middelburg and Lelystad through the good offices of the Rijkswaterstaat.

There were four designated metals of interest: copper, zinc, cadmium and mercury. The JMG indicated target concentrations of 1, 1, 0.1 and 0.02 micrograms per litre for the four metals respectively, and a salinity of 10 to 20 parts per thousand. The Western Scheldt estuary was recommended as the sampling site.

The coordinator suggested that the exercise would be of more value if two samples were collected. One would be of the required low salinity and the second of higher salinity, but from the same estuary. It was agreed that this would be done.

This exercise involves only the dissolved metal contents of the water, i.e. that portion passed by a 0.45 µm filter.

## SEAWATER COLLECTION AND SAMPLE PREPARATION

The sample collection and preparation was carried out by Alec Mykytiuk, Scott Wille and Shier Berman of NRCC and Phil Yeats from the Bedford Institute of Oceanography. All of these personnel have participated in previous collections of seawater for the preparation of NRCC reference materials for trace metals.

The equipment used for this exercise was essentially prepared and cleaned as described in a previous publication (Berman et. al., 1983). The water was peristaltically pumped aboard the M.V. Argus from about a 3 metre depth using silicone tubing, simultaneously filtered through 0.45 µm porosity acrylic copolymer filters (Gelman Sciences Inc.), acidified with high purity nitric acid and delivered into clean 50-litre polypropylene carboys.

A transect of the estuary from the Belgian border to a point about 9 kilometres ESE of Vlissingen was carried out on May 21, 1986. Samples were taken at 7 stations with salinities ranging from 8 to 28 parts per thousand. The samples were analyzed that same night at the Rijkswaterstaat laboratory in Middelburg and decisions were made as to where to collect the exercise samples consistent with the wishes of the JMG. The transect data will be published later.

Three hundred litres of water were collected on the morning of May 22 in the Western Scheidt estuary at 4 deg 2 min 12 sec E and 51 deg 23 min 24 sec N, about 7 km SSE of Hansveert and about 0.5 km off Walsoorden on the southern shore. Two hundred litres were acidified to pH 1.6 (1.5 ml acid per litre) and 100 litres were acidified with 400 ml of acid. The latter was to be the sample for mercury analysis, Sample C, and the former the desired JMG material, Sample B. Both have a salinity of 12 parts per thousand.

That afternoon, a two hundred litre sample was collected at 3 deg 42 min 35 sec E and 51 deg 24 min 37 sec N about 9 km ESE of Vlissingen approximately 2 km from the north shore. This water (Sample A), of salinity 28 parts per thousand, was filtered and acidified similarly to Sample B. It should be noted that this sample was collected well within the estuary.

The water was taken to the Lelystad laboratory where it was bottled in a clean area during the next four days. Each sample was equilibrated in a 200-litre clean polyethylene tank overnight and then peristaltically pumped into the respective bottles. Samples A and B were delivered into 2-litre polyethylene bottles and Sample C (for mercury) into 1-litre borosilicate glass bottles. Thus, 100 bottles of each sample were prepared.

The dissolved organic carbon contents were 1.03 mg/litre for Sample A and 2.46 mg/litre for Samples B and C.

#### SAMPLE DISTRIBUTION AND RECEIPT OF RESULTS

Seventy-eight sets of samples were sent to the avowed participants. Almost all of these were sent by air post from Lelystad during the week of May 25. A few samples were sent from Ottawa, Canada, at later dates in response to late requests for participation. Thirty-two of the sets were sent to JMP laboratories in accordance with a list supplied by the secretary of the Oslo and Paris Commissions and forty-six sets were sent to other ICES member country laboratories who asked to be included in the exercise.

The deadline for the receipt of results was set at September 30, 1986. This allowed at least four months for the completion of the work and submission of results. A reminder regarding the deadline was sent out in mid September. Twenty-three sets of results were received by September 30, and another nineteen sets by the time the final calculations were made in early November. The response of forty-two laboratories out of seventy-eight (54 percent) is very poor when compared with the 79 percent response for ICES 5/TM/SW. Approximately two-thirds of the JMP laboratories submitted results. This was somewhat better than the forty-six percent participation of the other laboratories but disappointing to the coordinator because he had been led to believe that participation was compulsory for the JMP laboratories. Nineteen laboratories sent neither a regret nor an excuse for not sending in results. These exercises are very costly in time and financial resources. The total cost to the National Research Council, the Rijkswaterstaat and the Commissions is in the region of 75,000 US dollars.

The participants were asked to carry out six replicate analyses of Samples A and B for the three compulsory metals, copper, zinc and cadmium and for any other metals they wished. Sample C was to be analyzed for mercury only, also using six replicate sub-samples. They were also asked to describe their procedures in detail, including calibration methods, limits of detection and whether they normally analyzed seawater for each of the metals. The participants were also warned that less than four replicate analyses would invalidate the results for any metal.

The large majority of respondents supplied the required information except that regarding their experience. But a few gave so little information about their procedures that it was impossible to adequately categorize them.

Unfortunately, a few sets of samples sent from Ottawa did not contain the instructions and two laboratories remitted results using only three replicate samples. These were accepted into the data sets since it seemed to be the only fair thing to do.

A number was assigned to each respondent on the receipt of results.

#### SAMPLE HOMOGENEITY

No extensive interbottle homogeneity testing was carried out. Our previous experience has shown no interbottle homogeneity problems with this method of sample preparation. The results of Lab 43A are from using four randomly chosen bottles for the six replicate samples. The precision is indistinguishable from that expected were all replicates taken from the same bottle.

One laboratory reported "the existence of a lot of suspended particles of considerable size" in Sample B. We have at times observed this in our bottled samples. It is due to polyethylene particles abraded from the bottle neck thread and has no effect on the trace metal contents.

#### RESULTS

All results, except a few very obvious outliers in some sets, were tabulated.

The number of significant figures in submitted data was reduced in many cases to a maximum of three for tabulation and computation. No set of analyses was accepted which contained less than four replicate results except as described above.

An evaluation has been made for the results for the four core metals of this study (copper, zinc, cadmium and mercury), and also for the two elements for which five or more sets of results were received (nickel and lead).

A  $t$  test at the 95 percent confidence level was applied to the mean of each set of values submitted (e.g. Liteanu and Rica, 1980). Means were successively rejected until a single set of results was obtained. This, of course, implies a normal distribution of results, which may not be a valid assumption. Indeed, the distribution in many cases was found to be not normal but the process allows for the elimination of gross outliers. An overall mean, standard deviation and relative standard deviation (RSD) were calculated for the remaining values. This mean and standard deviation are plotted on each graph. The small discrepancy sometimes apparent in the RSD is the result of rounding off significant figures in reporting the mean and standard deviation.

An attempt has been made to represent all the data received for the six metals on the individual graphs for each sample:

The range of results from each laboratory and their mean is plotted. A number of laboratories submitted more than one set of results for various elements in the samples. These have been plotted sequentially and the term "labs" in this section refers to the number of sets of results, not the number of individual participating laboratories.

A range continued by a forward arrow  $\rightarrow$  beyond the right margin of the graph indicates that the range exceeds this boundary.

A forward arrow  $\rightarrow$  beginning at the right margin indicates that all values submitted exceed the boundary.

A backward arrow  $\leftarrow$  at the right margin indicates results less than a limit of detection beyond the boundary.

A backward arrow  $\leftarrow$  within the graph indicates results less than the marked limit of detection.

A backward arrow  $\leftarrow$  from the left margin indicates results less than the boundary value or a limit of detection which is less than this boundary.

Means marked by a "m" are those that have been retained after the  $t$  rejection test and are incorporated into the overall result. Means marked by a "+" were rejected.

The results of the various calculations are beside each graph. The total range for all values reported is presented. A mean has been calculated for all quantitative results.

The final means, referred to as the excluded means, are the consensus values for the concentrations of the five metals in Samples A and B and of mercury in Sample C. However, they probably are good estimates of the real concentrations only for copper, zinc and nickel. The standard deviations are the spread of results representing this particular population of JMP and ICES laboratories after outlying means have been rejected.

There is always a problem about what to do with "less than" results. They are rejected if they are obviously wrong. For example, if the result is given as  $<0.1$  when the true value is 0.3. However, if it is reasonable (e.g.  $<0.1$  when the value is 0.03) it is accepted since good information has been given.

They are also rejected when the "less than" number given is more than five times the concentration of the metal in samples. It is assumed in that case that the procedure used has not adequate sensitivity for the analysis at hand.

An effort was made to assess the accuracy of the results. This is generally very difficult but in this exercise there is a benchmark for five of the six metals.

The NRCC (Lab 43B) has the capability to perform stable isotope dilution mass spectrometry analysis of seawater for nickel, copper, zinc, cadmium and lead. The analysis for mercury is also possible but the mercury concentration was too low for reliable estimation with the procedure usually used. This method, which involves the equilibration of known amounts of stable isotopes of the analyte metals with the sample and the measurement of the resulting isotope ratios, is essentially an absolute method of analysis because no standards are required. Also, because isotopes of an element are chemically indistinguishable, quantitative separations and concentrations are not necessary. It is essential, however, that isotopic equilibrium be established and that adequate blank evaluations be performed. The main errors are in the accuracy of measurement of the isotope ratios and the estimation of the blank. The method used in this case was isotope dilution inductively coupled plasma mass spectrometry (IDICPMS) (McLaren *et al.*, 1985).

A mean and standard deviation were calculated from all results within plus or minus one-third of the IDICPMS value. This is an arbitrary choice but allows for a range of accepted values which can differ by a factor of two from the lowest to the highest. It may be assumed that the results from a number of laboratories would tend to group about the "true value". This may also yield a significant improvement in precision through the elimination of random high and low values.

## DISCUSSION and CONCLUSIONS

In the following discussion the terms "accuracy" and "success" are used. They refer to participants' results which have fallen within about a factor of two standard deviations of the excluded mean calculated from all results within plus or minus one-third of the isotope dilution inductively coupled plasma mass spectrometric (IDICPMS) values for copper, zinc, cadmium, nickel and lead. The excluded mean for mercury was calculated as described on page 15. Thus, "accurate" determinations vary from plus or minus sixteen percent of the copper mean for Sample B to plus or minus thirty-eight percent of the zinc mean for Sample A for the four core metals. Accuracy, then, has been defined relative to the group performance and not in absolute terms.

An organization such as the JMG must set its own criteria for acceptable accuracy and laboratory precision for the various metals at various concentrations. An examination of the preceding performance tables will indicate which laboratories are capable of meeting these criteria.

While it is felt that good estimates of the concentrations of six metals in the three samples have been achieved and that the accuracy of the responding laboratories can be evaluated, caution should be observed in evaluating laboratory precision from a one-time test. The true intralaboratory precision can be known only through long term experience and adequate quality control procedures instituted in the individual laboratories.

Laboratory results were not weighted with respect to their precision in calculating the means as there is no reason to assume that there is a good correlation between intralaboratory precision and accuracy for these sets of results.

It is unfortunate that the participation rate of the JMP and, indeed, all the laboratories was not higher. The statistical base for the study has been eroded and the JMG will not be able to make as general a set of conclusions as they may have wished concerning the total population of JMP laboratories. Furthermore, the cost of these exercises is very high both in monetary terms and personal resources of the organizers. It may be wise in the future to levy a formidable participation fee, refundable on the submission of results. It might even be wiser to implement a system of compulsory participation in these studies.

Only two JMP laboratories (Labs 6 and 11) and one other (Lab 43) demonstrated the ability to accurately analyze the three estuarine water samples for the four core metals copper, zinc, cadmium and mercury. Four more laboratories (Labs 5, 23, 24 and 42), none of them JMP, were able to analyze Samples A and B for copper, zinc and cadmium but either did not submit mercury values for Sample C (Labs 23, 24 and 42) or the mercury result was rejected (Lab 5). The monitoring abilities of the JMP laboratories (and the other laboratories) for these metals in estuarine waters are clearly greatly limited.

Only eleven (Labs 1, 3, 6, 10, 11, 15, 19, 20, 26, 28, and 37), four (Labs 6, 10, 11, and 20) and seven (Labs 1, 3, 6, 11, 19, 21, and 26) JMP laboratories could successfully analyze both Samples A and B for either copper or zinc or cadmium respectively. Copper capability is the highest (fifty-two percent) while zinc and cadmium capabilities are nineteen and thirty-three percent respectively. If the three laboratories that apparently only analyze for mercury are omitted from the calculations these numbers rise to sixty-one, twenty-two and thirty-nine percent respectively.

It may be argued that Sample A with twenty-eight parts per thousand salinity and corresponding lower trace metal contents is not an "estuarine water" and was beyond the salinity limits requested by the JMG. The sample was, however, collected well within the Western Scheldt estuary in an area normally monitored as part of a national monitoring program. It is felt that both samples are representative of waters analyzed in many national programs.

In any case, there was no apparent difference in the capabilities of the laboratories to analyze both samples for copper and cadmium. Zinc, however, provided significantly greater difficulties in Sample A.

The use of gold to collect and concentrate mercury in the cold vapour atomic absorption spectrometry procedure is clearly advantageous. Those laboratories which do not use this procedure are advised to consider its adoption. However, use of this technique does not guarantee success as was demonstrated by the five rejected gold method results.

The three laboratories (all JMP) that only analyzed for mercury (Labs 14, 30 and 35) did well. It might be assumed that they are experienced mercury analysts. All rejected results for mercury were high which suggests the possibility that contamination and inadequate measurement of the procedural blank may be the greatest cause of poor performance for this metal.

Almost all the laboratories submitted quantitative results for copper. Twenty-one of the thirty-nine laboratories (fifty-four percent) submitting sets of values had accurate results for copper in both samples. This is the highest rate of success of the four core metals. The JMP laboratory rate is sixty-five percent.

The choice of analytical procedure for copper is not as clear cut as for mercury. Statistically, of the two common procedures, chelation/solvent extraction/AAS methods were superior to ASV with a success rate of eighty percent for the former versus sixty-one percent for the latter. But if six out of ten laboratories can carry out a procedure successfully then one wonders if the fault in the remaining laboratories is with the methodology or with the analysts and/or with the conditions under which they must perform.

It is obvious that both techniques are capable of producing good results for copper. There were no apparent differences in performance between labs that evaporated the extracting solvent or back extracted into an acid solution and those labs that determined the metal directly in the solvent.

There are too few data regarding the other procedures used to make general comments. It is surprising, however, that such poor performance was achieved using direct GFAAS (fifty percent) considering the relatively high concentrations of copper.

About seventy-five percent of the laboratories submitted quantitative results for zinc. Four laboratories indicated insufficient limits of detection for zinc in Sample A (Labs 18, 22, 28 and 39). This is hard to understand in view of the high sensitivities of atomic absorption spectroscopy methods, even FAAS, for this metal. Only eleven of the thirty-four laboratories (thirty-two percent) submitting values had accurate results for zinc in both samples. This is the poorest rate of success of the four metals. The JMP laboratory rate is about the same as the overall group at thirty-six percent.

There is a marked difference in performance for zinc between Sample A and Sample B. The latter yielded a normal distribution of results after the elimination of six outlying values (four low and two high). The former gave an abnormal distribution skewed to the high side even after the rejection of three very high values. When tested against the mean of all those labs within plus or minus one-third of the IDICPMS result another ten values were rejected (four low and six high).

An examination of the methodologies showed a clear superiority of performance of chelation/solvent extraction over ASV. The success rate of the former was sixty-four percent and eighty-three percent for the two samples respectively compared with an ASV success rate of only twenty-nine and sixty percent. It is apparently much more difficult to analyze seawater for zinc than for copper by ASV even when they are present at equivalent concentrations. Because sensitivity is not a factor at these concentrations for ASV the problem probably lies with

interferences introduced by the dissolved organic carbon content and/or the inability to control and adequately estimate the laboratory procedural blank.

Five laboratories analyzed the samples directly for zinc using GFAAS. The success rate is fifty percent.

Zinc is a ubiquitous metal. Great care must be taken in order to prevent sample contamination.

Almost all the laboratories submitted quantitative results for cadmium. Three laboratories indicated insufficient limits of detection for cadmium in one or both of the samples (Labs 20, 22 and 28). The latter two also had indicated insufficient sensitivity for zinc in Sample A. Eighteen of the thirty-eight laboratories (forty-seven percent) submitting quantitative values had accurate results for cadmium in both samples. The JMP laboratory rate was forty-seven percent. Both samples showed abnormal distributions skewed to the high side even after the rejection of outliers. Tests against the means of all those labs within plus or minus one-third of the IDICPMS results resulted in nineteen rejected values for each sample.

In the case of cadmium we see a reversal in the success rates of the two most common procedures. The rate for chelation/solvent extraction is forty-eight percent and that for ASV is fifty-nine percent. There is no difference in performance between samples.

ASV is as successful for cadmium as it is for copper in spite of the large differences in concentrations of the two metals. The chelation/solvent extraction success rate has dropped considerably for this metal. Only one-third of all direct GFAAS results are acceptable.

Eighty-seven percent of the rejected cadmium values were high. This may indicate that contamination and/or the inability to control and adequately estimate the laboratory procedural blank are the main factors which degenerate performance.

There were sufficient data submitted for both nickel and lead to warrant their inclusion in this study.

Almost forty percent of the laboratories submitted results for nickel. Eleven of these were JMP laboratories. Thirteen of the sixteen laboratories (eighty-one percent) could produce accurate values for nickel in both samples. This is by far the highest success rate of all the six metals studied.

Almost all the laboratories used a chelation/separation method. There were no ASV results. The high rate of success is apparently due to relatively high concentrations of nickel in the sample combined with the historically good behaviour of nickel in chelation/solvent extraction systems.

A surprisingly large number of laboratories (twenty-nine) submitted results for lead even though six were "less than" values. However, only seven laboratories (twenty-four percent) were able to produce accurate values for lead in both samples. Three of these were JMP laboratories.

Performances of the two common methods are equivalent at a forty percent success rate. All but one of the twenty-seven rejected results are high. The capability to analyze seawater for lead, even at concentrations up to two orders of magnitude higher than in open ocean waters, apparently still lies in the hands of a small group of analysts.

Generally, the results of this study are rather disappointing as it shows no improvements for copper, zinc and cadmium over that achieved in 1982 in the ICES 5/TM/SW exercise, especially when the differences in metal concentrations are taken into account. The analysis for mercury



was not a part of that study. No attempt was made to correlate individual performances of JMP laboratories which took part in ICES S/TM/SW with their performances in this exercise.

This laboratory (Lab 43) has, at one time or another, successfully employed all the procedures reported by the respondents, except TXRF, for the determination of trace metals in seawater. It is seen from this report that our favorite procedure involves the use of an immobilized ligand to separate and concentrate the trace metals from the seawater matrix. Laboratory 11, for example, uses chelation/solvent extraction with equal efficacy for the five metals. Laboratory 6 successfully employs a totally different variant of that method. There is obviously more than one way to carry out a successful analysis.

Electroanalytical methods such as ASV and direct determinations using GFAAS offer seemingly simpler methodologies when compared to separation and concentration procedures. However, this laboratory has become very wary of the ASV methods when the sample contains substantial quantities of dissolved organic carbon as is commonly found in estuarine waters. Interferences can sometimes be reduced by irradiation of the sample with ultraviolet light prior to the analysis but our experience has often shown increased and erratic blanks, especially for zinc, even when the irradiation cell, made of high purity quartz, has been meticulously cleaned. It is also possible that laboratories are not differentiating between the choice of the mercury drop and mercury film electrodes for specific analyses.

Direct determination by GFAAS is often not sensitive enough for seawaters. For example, the cadmium concentration in Sample A is only about twice the limit of detection for this metal by GFAAS making quantitative analysis a rather untenable situation. The direct analysis of lead should not even be attempted for these samples.

Based on experiences in other intercomparison exercises and suspicions in this one, it may be assumed that at least ten percent of poor results are due to arithmetical errors rather than analytical performance. Poor laboratory standards have also often been found to be the source of major errors. Clean working areas are mandatory.

If laboratories are not yet doing so, it is imperative that an adequate quality assurance program be implemented. It is beyond the intent of this report to discuss quality assurance which is a topic unto itself, but unless a laboratory can demonstrate by the use of reference materials which are now available and some very elementary statistics that its analyses are under statistical control then there is probably a considerable waste of resources and meaningless numbers are entering the data banks.

#### ACKNOWLEDGEMENTS

The authors gratefully acknowledge the efforts of Alec Mykytiuk and Scott Willie of this laboratory as well as Phil Yeats of the Bedford Institute of Oceanography who prepared the equipment, collected the water and bottled the samples. Arne Jensen from the Marine Pollution Laboratory in Charlottenlund and Wim Cofino from the Rijkswaterstaat in The Hague planned the exercise. Cofino also arranged for all the facilities necessary in the Netherlands including the M.V. Argus and ensured the smooth operation of the land and sea activities. We are also indebted to A. van den Akker and W. Scheurs for permission to use their laboratories in Lelystad and Middelburg respectively.

Special thanks go to Albert Westdijk and the crew of the Argus whose cooperation and congeniality made the sample collection possible and a pleasure.

The isotope dilution inductively coupled plasma mass spectrometry analyses were carried out by Diane Beauchemin, Alec Mykytiuk and Jim McLaren from this laboratory. The dissolved organic carbon was determined by Peter Wangersky and Regine Maass of Dalhousie University in Halifax.