Analysis of Interlaboratory Performance in the Determination of Total Selenium in Water

Peter J. Steinhoff,1 Brian W. Smith,2 Dewey W. Warner,1 and Gregory Möller1,3
University of Idaho, Moscow, ID 83844-2203

This study explored the performance of experienced laboratories in the analysis for total selenium in water by a variety of analytical methods. The goal of the study was to examine intra- and interlaboratory variability. Replicates (n = 7) of 7 sample types that included a reference material of known Se concentration, natural waters, and treated wastewaters were submitted to 7 laboratories with prequalified Se analytical experience. Results of the study indicated wide ranges in minimum and maximum results, distinct differences in laboratory precision, and routine reporting of numerical results below statistical limits of quantitation. Hydride generation as a sample introduction technique demonstrated superior performance. In general, the study supports a caution advisory about using low-level Se data, especially results lower than about 10 µg Se/L, without quantitating the statistical uncertainty of the data. Because this study used data from samples that were submitted in bulk to participating laboratories prequalified for Se analytical expertise and experience, it can be considered a best-case demonstration of performance.

The observation of selenium toxicosis in aquatic ecosystems and laboratory studies has led to the establishment of acute and chronic aquatic biota criteria for Se in water (1). The chronic exposure criterion of 5 mg Se/L challenges laboratories providing data to support the risk assessment of biota in selenium-affected ecosystems. This need results from the fact that regulatory methods for the laboratory analysis of Se in water and wastewater often have limited performance for low-level determinations (2), especially in complex samples that contain potential interferences such as heavy metals or high levels of total dissolved solids. Coupled with the potential presence of a variety of Se-containing compounds (both inorganic and organic) in multiple oxidation states for environmental Se, the precise and accurate laboratory determination of low microgram-per-liter levels in water is considerably more challenging than the determination of many other inorganic analytes of interest. For data users in the scientific and regulatory arenas, an accurate estimate of data quality, including statistical uncertainty, is required for risk assessment and enforcement. In addition, an understanding of the biogeochemical behavior of Se requires a high level of confidence in the data used to support scientific conclusions.

In this study, we examined the analytical performance statistics of a pool of laboratories to gauge the reliability of Se data. The study design targeted laboratories that have significant experience in the determination of total Se in water. Laboratories were prequalified for expertise in Se determination by external references or prior participation in large-scale Se monitoring or assessment projects. In this study, we statistically examined the performance of this pool of laboratories in an effort to describe both intra- and interlaboratory variability.

Background of Selenium Determination

Total selenium in water is determined routinely by several approaches (3). Digestion of complex matrices and manipulation of the Se oxidation state are typically accomplished by using a variety of acids, bases, or other oxidizing and reducing agents. Various digestion and preconcentration approaches have been used to modify the sample matrix and Se oxidation states (4–6).

In graphite furnace atomic absorption analysis (GFAA), Se is electrothermally atomized and absorption of 196 nm light is measured. Hydride generation atomic absorption spectroscopy (HGAA) uses the reaction of the selenite species with sodium borohydride to produce Se hydride, followed by gas–liquid separation to volatilize Se into a high-temperature quartz cell in an absorption pathway (6–8). GFAA for Se is difficult with samples containing high levels of dissolved solids; therefore, matrix modifiers are typically used to volatilize salts. Interferences in HGAA include the selenide-forming elements, such as copper (9).

Selenium is also routinely determined by inductively coupled argon plasma using atomic emission spectroscopy (ICPAS) and mass selective detection (ICPMS; 10–12). Because only about 5% of the aspirated sample can be introduced into the plasma at typical rates of flow injection nebulization,
analytical sensitivity can be significantly enhanced by modified delivery of Se into the plasma, such as volatile SeH₂ generation. ICPMS follows the argon plasma atomization and ionization of Se by mass selective detection. Atomic and polyatomic mass-charge interferences can require compensation by mass resolution and computer-based interference management in determinations of low levels of Se.

Atomic fluorescence (AF) and molecular fluorescence approaches have been successfully applied to Se determination (3, 13, 14). In a typical approach for Se, hydride generation and gas-liquid separation are used to introduce Se hydride into a flame where it is atomized and determined in fluorescence orientation. Complexation of Se with fluorophores such as naphthalene-2,3-diamine-Se(IV) is the basis of the fluorimetric determination of Se (3, 4, 14). The Se species must be converted to Se(IV) and transferred to an organic solvent for analysis. Sources of difficulties in the fluorimetric determination of Se include sample analysis time and degradation of the fluorescent complex by about 3% per week in normal light.

Neutron activation analysis (NAA) consists of bombarding a sample with neutrons in a nuclear reactor to change the original isotopes to radioactive isotopes of Se (15, 16). After this exposure, decay counting of the target radioactive isotopes is used to determine how much of the analyte is present. For Se, the long half-life of the decay process can allow for highly sensitive analysis by delaying counting for a period of 10–14 days. This delay ensures that similar, potentially interfering, energy emissions from elements with shorter half-lives are reduced during the counting stage of the analysis.

Table 1. Summary of interlaboratory data from analyses of water samples for total selenium content^a

<table>
<thead>
<tr>
<th>Sample</th>
<th>Mean range, μg Se/L</th>
<th>Minimum, μg Se/L</th>
<th>Maximum, μg Se/L</th>
<th>MDL range, μg Se/L^b</th>
<th>LOQ range, μg Se/L^b</th>
</tr>
</thead>
<tbody>
<tr>
<td>FS5</td>
<td>4.5–7.4</td>
<td>2.9</td>
<td>10.1</td>
<td>0.3–5.8</td>
<td>1.1–19.3</td>
</tr>
<tr>
<td>FS50</td>
<td>40.5–58.1</td>
<td>38.8</td>
<td>65.7</td>
<td>2.3–15.2</td>
<td>7.6–50.6</td>
</tr>
<tr>
<td>Sta 10</td>
<td>3.4–6.0</td>
<td>2.2</td>
<td>8.8</td>
<td>0.2–6.4</td>
<td>0.6–21.3</td>
</tr>
<tr>
<td>NFA</td>
<td>1.8–4.3</td>
<td>0.2</td>
<td>6.0</td>
<td>0.1–7.6</td>
<td>0.5–25.4</td>
</tr>
<tr>
<td>NFB</td>
<td>1.7–2.8</td>
<td>–1.3</td>
<td>4.3</td>
<td>0.3–8.5</td>
<td>0.9–28.3</td>
</tr>
<tr>
<td>Drain</td>
<td>2.7–4.9</td>
<td>1.7</td>
<td>6.5</td>
<td>0.2–5.1</td>
<td>0.8–16.9</td>
</tr>
<tr>
<td>Outfall</td>
<td>10.0–17.2</td>
<td>8.2</td>
<td>18.8</td>
<td>0.8–12.2</td>
<td>2.6–40.6</td>
</tr>
</tbody>
</table>

^a Laboratory 7 data were not used because of lack of statistical weight (n = 3).

^b Method detection limit, MDL (3σ), and limit of quantitation, LOQ (10σ), were determined from each laboratory’s replicate data submission (n = 7).

Figure 1. Box plot of the selenium interlaboratory results for the 5 μg/L standard (n = 7). Circles represent the data extremes, the T bars represent the 10th and 90th percentiles of the data range, the top of the box is the 75 quartile, the bottom of the box is the 25 quartile, and the line in the box is the median of the data. L7 had insufficient replicates (n = 3) to allow full data analysis.
Typically, the heat of the conditions of reactor exposure is not compatible with liquid water. To accomplish NAA of water samples, the target analytes must be sorbed on a solid matrix, or the water must be removed by evaporation or freeze-drying.

**Detection and Quantitation of Selenium**

In analytical work, 3 concentration ranges are of interest. In the first region, the presence of the analyte is statistically indeterminate. In the second region, the analyte is present but cannot be accurately quantitated, and in the third region the analyte is quantitatable. There is some disagreement on the best approach for calculating the limits for these concentration ranges.

The U.S. Environmental Protection Agency (EPA; 1,17) defines method detection limit (MDL) as the minimum concentration of an analyte that can be identified, measured, and reported with 99% confidence (that the analyte concentration
is greater than zero) as determined from replicate analyses of a sample of a given matrix containing the analyte. Functionally, this value is the standard deviation (σ) of the analytical results for the samples multiplied by the Student’s t statistic for the degrees of freedom related to the number of samples. Conversely, the American Chemical Society (ACS) defines MDL by using either a sample or a blank and 3σ for the recommended MDL calculation (18). The MDLs calculated by the ACS definition are smaller than the MDLs calculated from the EPA definition if fewer than 8 samples are used and greater if more than 8 samples are used. Additionally, MDLs are somewhat instrument and matrix specific. Compliance monitoring is not instrument and matrix specific, and analysis may be performed at more than one laboratory. Maddalone et al. (19) proposed a compliance monitoring detection level (CMDL) based on a 99% confidence level and a standard deviation calculated directly from all relevant data instead of a pooled standard deviation from the laboratories involved.

Limit of quantitation (LOQ) and compliance monitoring quantitation level (CMQL) are intra- and interlaboratory computations that define a concentration level or amount of analyte above which measured levels are considered quantitatable. The LOQ is computed by using the same standard deviation that was used for computing the MDL. The ACS recommends a level of 10σ for the LOQ for a possible 30% error (10σ ± 3σ). Maddalone et al. (19) propose that the standard deviation at the quantitation level is significantly different from that at the CMDL level, and that a CMQL should

Figure 4. Box plot of the selenium interlaboratory results for the Outfall samples (n = 7). Plot is described in Figure 1 legend. L7 had insufficient replicates (n = 3) to allow full data analysis.

Figure 5. Comparison of mean selenium concentration results for the 5 μg/L standard with the calculated 3σ MDL and calculated 10σ LOQ for each study laboratory. L7 had fewer replicates (n = 3) than the other laboratories (n = 7).
be calculated by using a different standard deviation. They approximate \( \sigma \) from a plot of standard deviation versus concentration and multiply \( \sigma \) by the appropriate Student’s \( t \) statistic to give a 99% confidence limit. In this work, we use the 3\( \sigma \) MDL and 10\( \sigma \) LOQ approach suggested by the ACS (18).

**Experimental**

**Study Samples**

(a) **Environmental water and wastewater samples.**—Five bulk water samples, typical of the monitoring samples used in the estimation of environmental background and treated discharge water quality, were obtained from the Cominco–Alaska, Red Dog zinc mine site (Red Dog, AK): (1) **Station 10 (Sta10).**—Stream water near the confluence of Red Dog Creek and Ikalukrok Creek, creek water below a treated wastewater-mixing zone. (2) **North Fork A (NFA).**—North Fork of Red Dog Creek, natural creek water. (3) **North Fork B (NFB).**—North Fork of Red Dog Creek, natural creek water. (4) **Drain (Drain).**—Precipitation drainage from the mine site and waste piles, mixed acid rock drainage. (5) **Outfall (Outfall).**—High level of sulfate, high level of total dissolved solids, lime-treated mill, and tailings pond effluent, ca pH 9.

(b) **Selenium-fortified water samples.**—Fortified reference samples were prepared from Na\(_2\)SeO\(_3\) (Sigma No. S-1382; Sigma Chemical Co., St. Louis, MO) and Na\(_2\)SeO\(_4\) (Sigma No. S-0882) in 18 M\( \Omega \)-cm water. The 5 mg Se/L (FS5) and 50 mg Se/L (FS50) samples were prepared as a 50 + 50 mixture of Na\(_2\)SeO\(_3\) and Na\(_2\)SeO\(_4\) by using serial dilution.

**Sample Preparation and Submission**

All samples were bulk-filtered through glass fiber filter paper (Micro Filtration Systems, Dublin, CA; GF 75 grade) by using pressure filtration (Millipore Corp., Bedford, MA; Model YT 30 142 HW). All samples were brought to pH 2 with nitric acid and stored in bulk until they were transferred...
into 125 mL high-density polyethylene sample bottles. Each sample bottle was labeled with a unique randomized sample number, and the Se-fortified water samples were identified by the target concentration of the sample (5 or 50 µg Se/L). Each laboratory was assigned 7 replicates of each of the 7 water samples. Only 3 replicates were submitted for NAA. The study samples for each laboratory were randomly divided into 3 batch packages and shipped in coolers to the participating laboratories. The study samples were submitted to the laboratories with the request that the batches be analyzed on 3 different days by the most sensitive, analytical method for total Se in water that the laboratory routinely performs.

Participating Laboratories and Methods of Analysis

Laboratory 1 (L1) is a large scale, commercial contract laboratory that provides analytical laboratory support for governmental and industry clients. EPA Method 6010, Inductively Coupled Plasma Atomic Emission Spectroscopy, was used to analyze the samples for their Se content. Samples were injected directly into a high sensitivity “trace” ICPAES instrument.

Laboratory 2 (L2) is a large-scale, commercial contract laboratory that provides analytical laboratory support for governmental and industry clients. The sample was digested first in nitric acid and then in hydrogen peroxide as described in EPA ILMO4.0 (EPA Contract Laboratory Program Statement of Work, Multi-Media Multi-Concentration). EPA Method 7740, GFAA for Selenium, was used to determine the amount of Se in each sample.

Laboratory 3 (L3) is a large-scale, commercial, geochemical research and production analysis laboratory. Samples were oxidized with potassium persulfate and then reduced with 40% HCl solution. Analysis was by hydride generation atomic fluorescence spectroscopy. This method has been proposed as a draft EPA method (13).

Laboratory 4 (L4) is a large-scale, university research support laboratory. The sample was digested in a 3-step process. In the first step, the sample was concentrated by nitrogen

Table 2. Summary of data relative to calculated method detection limits (MDLs) and calculated limits of quantitation (LOQs) as reported by 7 laboratories

<table>
<thead>
<tr>
<th>Sample</th>
<th>Mean reported concentration range, µg Se/L (n = 7)</th>
<th>No. of labs reporting data below their 3σ calculated MDL</th>
<th>No. of labs reporting data below their 10σ calculated LOQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>FS5</td>
<td>4.1–7.4</td>
<td>1 (L7)</td>
<td>3 (L1, L2, L7)</td>
</tr>
<tr>
<td>FS50</td>
<td>37.6–58.1</td>
<td>0</td>
<td>1 (L7)</td>
</tr>
<tr>
<td>Sta 10</td>
<td>3.4–6.0</td>
<td>0</td>
<td>3 (L1, L2, L7)</td>
</tr>
<tr>
<td>NFA</td>
<td>1.6–4.3</td>
<td>2 (L1, L2)</td>
<td>5 (L1, L2, L5, L6, L7)</td>
</tr>
<tr>
<td>NFB</td>
<td>1.7–2.8</td>
<td>1 (L1)</td>
<td>5 (L1, L2, L5, L6, L7)</td>
</tr>
<tr>
<td>Drain</td>
<td>2.7–12.8</td>
<td>1 (L2)</td>
<td>3 (L2, L6, L7)</td>
</tr>
<tr>
<td>Outfall</td>
<td>10.0–17.2</td>
<td>0</td>
<td>3 (L1, L2, L7)</td>
</tr>
</tbody>
</table>

* Standard deviation (σ) was determined from the reported analytical results for 7 replicates of the 7 samples for each laboratory.
* Laboratory 6 provided analytical data, but indicated results were below the practical quantitation limit.
evaporation and heated with nitric acid. In the second step, the sample was boiled in a mixture of sulfuric and perchloric acids. Finally, the sample was reduced with HCl. The digested sample was analyzed by hydride generation ICPAES (8).

Laboratory 5 (L5) is a large-scale, university research support laboratory that used the same procedure as L4 except that it did not concentrate the sample with nitrogen evaporation.

Laboratory 6 (L6) is a large-scale, commercial contract laboratory that performs analytical laboratory support for governmental and industry clients. Determination of Se content was by HGAA in a method similar to EPA Method 7741, Selenium (AA, Gaseous Hydride). The sample was digested in hydrogen peroxide and NaOH at 95°C. After cooling and a helium purge, the samples were reheated and reduced with HCl.

Laboratory 7 (L7) is a university research nuclear reactor performing NAA. Samples were prepared by freeze-drying in cleaned polyethylene bags. The bags were then placed in a cylinder for neutron bombardment in the reactor. The samples were irradiated at a flux of $6.53 \times 10^{12} \text{n/cm}^2/\text{s}$ for 6 h. Samples were held for 14 days so that interfering element emissions would be reduced. After 14 days, samples were counted at the 136.0, 264.7, and 279.5 keV $^{75}\text{Se}$ gamma emissions. The 279.5 keV peak was used for Se quantitation. Sample counts were normalized to standard reference materials (water and urban particulate) from the National Institute of Standards and Technology (Gaithersburg, MD) for Se quantitation.

**Statistical Analyses**

Standard statistics and analysis of variance (ANOVA) using statistical software (SAS Institute, Inc., Cary, NC; Version 5) and spreadsheet software (Excel 5.0, Microsoft Corp., Redmond, WA) were used to examine the data in this study. The ACS methods (18) were used for calculating the MDLs (3σ) and LOQs (10σ). These calculations were applied to the replicate Se results in the study.

**Results and Discussion**

All the laboratories in the study submitted data. Standard statistical analysis and graphical analysis yielded results that could be used to interpret the performance of the laboratories in the pool. ANOVA using means analysis for significant differences was not successful because of the unacceptable range of the variances in the intralaboratory data.

**Summary Statistics**

Table 1 shows the summary statistics for the study data and includes the MDL and LOQ ranges calculated from each laboratory’s results. The range of means of the replicate analyses for the 5 μg/L Se reference sample was 58% of the target value, whereas that for the 50 μg/L Se reference sample was 33% of the target value. Maximum/minimum data for the study demonstrate single sample result variability across the pool.

**Box Plot Analysis**

Box plots of the data, showing the data extremes, medians, and distributions, are shown in Figures 1–4. In Figure 1, the results for the 5 μg Se/L reference sample illustrate a clustering of the medians of L2–L7. L1 shows a distinct positive bias. L3–L6 demonstrate good accuracy and precision for this sample set.

The NFA data box plots are shown in Figure 2. L2–L7 show consistent means, and L3 and L4 demonstrate very high precision. The Drain sample data plotted in Figure 3 show the increasing performance variability for samples containing larger quantities of analytical interferences. This sample contained higher concentrations of zinc and lead than were present in the other samples. L1 did not produce a quantitative result and reported <80 μg Se/L for the test result. An examination of the raw data for this analysis shows the presence of a significant positive interference in the results for the L1 and L7 methods of analysis. L2 results showed a wider range of values than L3–L6, which showed clustered data. L3 and L4 demonstrated the highest precision within this cluster. L2 showed a high bias, and L6 showed a low bias relative to the laboratories with consistent performance.

The data for the Outfall sample, which contained high levels of dissolved solids as potential interferences, are represented in Figure 4. L3–L7 had clustered medians. L1 showed a high bias relative to this cluster and L2 maintained a low bias. L3–L6 demonstrated good precision for this very complex sample. It is worth noting that this sample is used in determining discharge permit compliance for the mine and mill site involved in this study.

**Histogram Analysis**

The replicate data submitted by the laboratories allow for a working MDL and LOQ to be calculated for each sample for each laboratory. These calculated performance indicators can yield an estimate of the numerical significance of the data produced by these facilities, while at the same time allowing for qualitative estimation of the global capability of this pool of laboratories to determine total Se in water.

Bar graphs of the data means, MDLs, and LOQs are shown in Figures 5–8. Figure 5, which shows the results for the 5 μg Se/L sample, illustrates a clustering of means for L2–L7. L1 shows a distinct positive bias. L3–L6 demonstrate high precision and have mean concentrations greater than their LOQs.

The results of the NFA analyses are shown in Figure 6. Only L3 and L4 reported means above their LOQs, although all laboratories except L1 reported means consistent with the means reported by L3 and L4. Statistically, the means reported by L1 and L2 cannot be determined to be nonzero.

Mine water samples analyzed in this study included Drain and Outfall. Results from the Drain analyses are presented in Figure 7. The reported means have the greatest percentage spread in the study. L1 could not determine the presence of Se, and L7 reported an Se concentration that is significantly higher than those for the other laboratories. L2–L6 exhibit a clustering of means, although the mean reported by L2 cannot
be statistically determined to be nonzero. Figure 8 shows the results of the Outfall analyses. L3, L5, and L6 show a clustering of means as do L4 and L7, whereas L1 and L2 are study extremes. L3–L6 show a high level of precision, whereas L1 and L2 show a lack of precision.

**MDL and LOQ Performance Analysis**

Table 2 shows how each laboratory reported data relative to its calculated MDLs and LOQs. L7 always reported data below its LOQ, but only once reported it below its MDL; however, the small sample size \((n = 3)\) limits the accuracy of this performance estimate. L1 and L2 most frequently reported data below their MDLs. L3 and L4 always exceeded their MDLs and LOQs. Of all the laboratories reporting data below their 10\(\sigma\) LOQs, only L6 reported that the data were below the LOQ [practical quantitation limit (PQL)]. None of the laboratories reported data in the manner suggested by the data acquisition and data quality evaluation guidelines (18). The suggested method for reporting data >MDL and <LOQ is listing the numerical test result alongside the LOQ for the test sample type.

**Conclusions**

The results of this study indicate that there is a wide range of performance quality in the experienced analytical laboratories in this pool. The range of results for individual samples indicates the need for replicate sampling to determine the statistical uncertainty of the data from analyses for low concentration, total Se in water. The 4 laboratories using hydride generation as a sample introduction technique (L3–L6) consistently reported precise and clustered data, as compared with the other participants. These results indicate the relative robustness of hydride generation in analytical methods for the determination of Se.

In the conduct of this study, the normal variability of day-to-day performance was modeled with 3 randomized batches and a request to analyze on different days. This design element can be considered representative; however, it would typically understate variability for larger time frames of performance. In this approach, there was limited opportunity to ensure multiday data generation in true blind batch fashion. Additionally, the study design was positive performance biased as a result of the prequalification requirements. Thus, the study design can be regarded as reflecting near best-case variability.

If we assume that this pool of laboratories is representative, we can speculate on the generic performance of laboratories experienced with analyses for total Se in water and the quality of the analytical data. The LOQ study mean calculated without the NAA results, without the L1–Drain performance fail-

**References**

2. South Dakota Department of Environmental and Natural Resources (1996) *Assessment of the Surface Water Quality and the Aquatic Ecosystem of the Northern Black Hills*, Pierre, SD