Performances of analytical methods for freshwater analysis assessed through intercomparison exercises. II. Major ions

Aldo MARCHETTO*, Rosario MOSELLO, Gabriele TARTARI, Michele BIANCHI1), Helga GEISS1), Giorgio SERRINI1), Gianna SERRINI-LANZA1) and Herbert MUNTAU1)

CNR Istituto Italiano di Idrobiologia, L.go V. Tonolli 50, 28922 Verbania Pallanza, Italy
1)EU Joint Research Centre, Environment Institute, 21020 Ispra (VA), Italy
*e-mail corresponding author: a.marchetto@iii.to.cnr.it

ABSTRACT

From 1991 to 1997, nine intercomparison exercises on rainwater and freshwater samples were held in the framework of an activity connecting laboratories participating in different projects (Environmental studies in the Mediterranean basin, Italian network for the study of wet deposition, Acidification of mountain lakes, Limnological studies of Lake Léman, Monitoring of atmospheric deposition in forests). The number of participants increased from 72 in 1991 to 177 in 1997. Among the measured variables (pH, conductivity, alkalinity, major ions and nutrients), Ca, Mg, Na, K, ammonium, sulphate, chloride and nitrate are considered in this paper. The methods commonly used by the laboratories involved in the exercises are compared for precision, and an estimate of the performance of each method at different concentrations is made. The results show that some of the participating laboratories use analytical methods which are not reliable for the concentration values occurring in freshwater and/or in rainwater. However, outliers are also produced by laboratories using well-performing methods, indicating the importance of introducing analytical control procedures.

Key words: intercomparison, freshwater, rainwater

1. INTRODUCTION

From 1991 to 1997, nine intercomparison exercises were held within the project «Analytical Quality Control and Assessment Studies in the Mediterranean Basin (AQUACON-MedBas)», part of the «Protection of the Environment» program, carried out by the Environment Institute of the Joint Research Centre (JRC-EI), in collaboration with the Istituto Italiano di Idrobiologia of the Italian National Research Council (CNR-III). The program, developed by the EU-member countries belonging to the Mediterranean area in close partnership with the Environment Institute, is aimed at the identification, quantification and reduction of random and systematic errors associated with the most important branches of environmental analysis.

Seven yearly exercises focussed on rainwater analysis, and two exercises on freshwater analysis. At their request, other research teams working in the environmental field joined the exercises, which eventually involved the laboratories participating in the following projects: Analytical Quality Control and Assessment Studies in the Mediterranean Basin (AQUACON), Acidification of Mountain Lakes: Palaeolimnology and Ecology (AL:PE), Mountain Lake Research (MOLAR), International Commission for the Protection of Lake Léman, Italian network for the study of wet deposition chemistry (RIDEP), UN-ECE International Co-operative Program on Assessment and Monitoring of Air Pollution Effects on Forests (ICP-Forests). A group of South American laboratories involved in freshwater research also asked to be included in the exercise. The list of participating laboratories and the whole set of results are reported elsewhere (Mosello et al. 1992, 1993, 1994, 1995, 1996, 1997a, 1997b, 1998a, 1998b). Having discussed alkalinity measurements (Marchetto et al. 1997), in this paper we use results obtained for major ions (Ca, Mg, Na, K, ammonium, chloride, sulphate and nitrate) by the participating laboratories to compare the relative performance of different analytical methods.

2. METHODS

2.1. Sample preparation and evaluation

For each exercise, four solutions were prepared at the EJ-IRC; two of them were intended for the measurement of pH, conductivity and major ion concentrations, while the other two solutions were specifically prepared for alkalinity (not discussed here) and nutrients (nitrogen, phosphorus and reactive silica), in the case of rainwater and freshwater exercises, respectively. Each sample was prepared using water of the highest quality (Nanopure U.W.S. Barnstead, conductivity below 0.006 mS m⁻¹) and the purest chemicals available. Target concentrations were selected to be comparable with those found in atmospheric deposition and freshwater in southern Europe. The carefully weighed chemicals were dissolved and water added to make up the stock solution (1 litre), which was then analysed to check the correctness of the envisaged analyte concentrations. The master solution was then diluted, using Nanopure water, to
about 20 l in a 100 l polyethylene container, previously conditioned for two weeks with Nanopure water. Suprapure HCl was then added and made up to a total of about 100 l. Bottling was done by hand, the previously conditioned (two weeks with Nanopure water) 500-ml polypropylene bottles being rinsed with the samples and then filled up to the top. Samples were mailed to the participating laboratories, and the stability of the solutions was checked by the two organising laboratories by analysing samples kept in the dark at room temperature over the period allowed for the exercise (Mosello 1992, 1993, 1994, 1995, 1996, 1997a, 1997b, 1998a, 1998b). The stability of the solutions for the duration of the exercise was demonstrated by the results of these analyses.

An estimation of the variance of each concentration was obtained by measuring it in ten randomly selected bottles for each sample. All the analyses were performed in one laboratory by the same analyst using the same analytical method. The variance due to the analytical method was then estimated by repeating the measurement ten times on the same bottle. Heterogeneity in the solutions was then estimated as the square root of the difference of the variances of samples and methods; this proved to be generally lower than 1.5%, and even lower than 1% in most cases. In the case of the 1997 freshwater samples, heterogeneity of between 1.5% and 3% was found in the case of ammonium, the concentration of which was lower than 0.1 mg-N l⁻¹.

The participating laboratories were requested to perform a single analysis for each sample using their preferred method and to document in detail the analytical procedures used.

2.2. Data elaboration

Target values (Tab. 1) were calculated as the mean of the values obtained by the organising laboratories, using ion chromatography for the anions, the indophenol blue spectrophotometric method for ammonium and ion chromatography and atomic absorption spectrophotometry for the major cations.

Outliers were detected as follows (Hovind 1989): data out of the range of ±50% of the median value were first identified as outliers. Then the mean and standard deviation (σ) of the remaining data were calculated and data out of the range ±3σ from the mean were iteratively detected as outliers.

Outlier accommodation (e.g., Huber 1984) was used to obtain estimates of averages and standard deviations slightly affected by the presence of a high number of outlying results. In this paper we used the iterative technique known as H15, which begins by assigning to the estimated robust mean (m₀) the median of the sample values (xᵢ) and to the estimated robust σ (s₀) the median of the quantities |xᵢ - m₀|/0.6745. Then, at each n-th iteration, all values higher than mₙ₋₁ + Fₛₙ₋₁ or lower than mₙ₋₁ - Fₛₙ₋₁ are replaced by the pseudo-values mₙ₋₁ + Fₛₙ₋₁ and mₙ₋₁ - Fₛₙ₋₁, respectively, while the pseudo-values for the remaining values are the values themselves. The new estimate of the robust mean mₙ will be the mean of the pseudo-values, while the new estimate of the robust σ (sₙ) will be their σ divided by the square root of a constant (E), to compensate for the downweighting of the extreme values. We assumed a value of 1.5 for c, and consequently a value of 0.736 for b (Analytical Method Committee 1989). The estimated values on the basis of analyses performed by the reference laboratories.

<table>
<thead>
<tr>
<th>Exercise</th>
<th>SO₄²⁻ mg l⁻¹</th>
<th>N–NO₃⁻ mg l⁻¹</th>
<th>Cl⁻ mg l⁻¹</th>
<th>Ca²⁺ mg l⁻¹</th>
<th>Mg²⁺ mg l⁻¹</th>
<th>Na⁺ mg l⁻¹</th>
<th>K⁺ mg l⁻¹</th>
<th>N–NH₄⁺ mg l⁻¹</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rainwater exercises</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1991A</td>
<td>2.4</td>
<td>0.11</td>
<td>0.18</td>
<td>0.21</td>
<td>0.10</td>
<td>0.17</td>
<td>0.23</td>
<td>0.32</td>
</tr>
<tr>
<td>1991B</td>
<td>11.7</td>
<td>0.79</td>
<td>2.04</td>
<td>0.81</td>
<td>0.21</td>
<td>1.20</td>
<td>0.32</td>
<td>1.26</td>
</tr>
<tr>
<td>1992A</td>
<td>2.6</td>
<td>0.13</td>
<td>0.16</td>
<td>0.20</td>
<td>0.10</td>
<td>0.20</td>
<td>0.29</td>
<td>0.38</td>
</tr>
<tr>
<td>1992B</td>
<td>12.0</td>
<td>1.04</td>
<td>1.19</td>
<td>1.26</td>
<td>0.32</td>
<td>1.48</td>
<td>0.43</td>
<td>1.17</td>
</tr>
<tr>
<td>1993A</td>
<td>3.5</td>
<td>0.12</td>
<td>0.87</td>
<td>0.36</td>
<td>0.14</td>
<td>0.17</td>
<td>0.20</td>
<td>0.38</td>
</tr>
<tr>
<td>1993B</td>
<td>14.5</td>
<td>0.97</td>
<td>1.91</td>
<td>1.08</td>
<td>0.34</td>
<td>0.70</td>
<td>0.45</td>
<td>1.40</td>
</tr>
<tr>
<td>1994A</td>
<td>4.3</td>
<td>0.15</td>
<td>0.34</td>
<td>0.29</td>
<td>0.09</td>
<td>0.20</td>
<td>0.25</td>
<td>0.41</td>
</tr>
<tr>
<td>1994B</td>
<td>13.8</td>
<td>1.16</td>
<td>1.95</td>
<td>1.32</td>
<td>0.29</td>
<td>1.64</td>
<td>0.45</td>
<td>1.15</td>
</tr>
<tr>
<td>1995A</td>
<td>2.7</td>
<td>0.15</td>
<td>0.93</td>
<td>0.20</td>
<td>0.10</td>
<td>0.21</td>
<td>0.24</td>
<td>0.39</td>
</tr>
<tr>
<td>1995B</td>
<td>11.0</td>
<td>1.05</td>
<td>3.40</td>
<td>1.27</td>
<td>0.30</td>
<td>1.48</td>
<td>0.40</td>
<td>1.24</td>
</tr>
<tr>
<td>1996A</td>
<td>3.9</td>
<td>0.17</td>
<td>1.18</td>
<td>0.24</td>
<td>0.11</td>
<td>0.20</td>
<td>0.31</td>
<td>0.40</td>
</tr>
<tr>
<td>1996B</td>
<td>14.6</td>
<td>1.10</td>
<td>3.35</td>
<td>1.48</td>
<td>0.40</td>
<td>1.53</td>
<td>0.42</td>
<td>1.30</td>
</tr>
<tr>
<td>1997A</td>
<td>11.7</td>
<td>1.23</td>
<td>3.15</td>
<td>1.25</td>
<td>0.50</td>
<td>1.73</td>
<td>0.50</td>
<td>1.35</td>
</tr>
<tr>
<td>1997B</td>
<td>2.7</td>
<td>0.16</td>
<td>0.61</td>
<td>0.20</td>
<td>0.13</td>
<td>0.20</td>
<td>0.40</td>
<td>0.40</td>
</tr>
<tr>
<td>Freshwater exercises</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1996A</td>
<td>13.5</td>
<td>0.31</td>
<td>3.9</td>
<td>18.8</td>
<td>2.7</td>
<td>3.2</td>
<td>2.1</td>
<td>-</td>
</tr>
<tr>
<td>1996B</td>
<td>62</td>
<td>0.63</td>
<td>8.3</td>
<td>41.5</td>
<td>7.0</td>
<td>7.1</td>
<td>3.5</td>
<td>-</td>
</tr>
<tr>
<td>1996C</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1996D</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-0.40</td>
<td>-</td>
</tr>
<tr>
<td>1997A</td>
<td>23.0</td>
<td>0.47</td>
<td>0.50</td>
<td>24.0</td>
<td>2.3</td>
<td>2.2</td>
<td>2.0</td>
<td>-</td>
</tr>
<tr>
<td>1997B</td>
<td>67</td>
<td>0.76</td>
<td>1.30</td>
<td>41.6</td>
<td>5.1</td>
<td>4.1</td>
<td>2.4</td>
<td>-</td>
</tr>
<tr>
<td>1997C</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-0.035</td>
<td>-</td>
</tr>
<tr>
<td>1997D</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-0.090</td>
<td>-</td>
</tr>
</tbody>
</table>
parameters rapidly converge to the robust mean and standard deviation.

3. RESULTS AND DISCUSSION

The relative standard deviation of the results is generally very high (Fig. 1) and the distributions of the results skewed toward high values, because of the presence of outlying results (Fig. 2). Outlier rejection would greatly reduce relative $\sigma$ (Fig. 3), but methods with a high percentage of rejected outliers (e.g., silver nitrate titration for chloride and EDTA titration for Ca and Mg) will then seem to perform better than the most reliable methods, because the rejection procedures will reject all results, apart from the few that happen to be close to the target.

Robust statistics were then used to compare the standard deviation of the results obtained for each variable using different analytical methods.

Methods used by more than five laboratories in any exercise are listed in Table 2 and compared in Figures 4 through 7. The most commonly used methods were ion chromatography (IC) for anions, the indophenol spectrophotometric method for ammonium and atomic absorption spectrometry (AAS) for major cations. Between 1991 and 1997 the use of IC and inductively coupled plasma spectrometry (ICP) for base cation analysis was increasing, while the number of laboratories using atomic emission spectrometry (AES) for the analysis of Na and K was declining (Tab. 2).

Averages (either parametric or robust) of the values obtained using different analytical methods are never significantly different from each other (Student’s $t$-test). Some methods (e.g. EDTA titration for Mg) greatly overestimate concentrations, but the relative $\sigma$ (either parametric or robust) of the results obtained using these methods are so high that no significant difference between means is detected by the $t$-test. Each method was compared in terms of the robust $\sigma$ of the results to the most used one, using the Snedecor’s $F$-test at 1% probability level, and the results are summarised in the plots (Figs 4-7). However, considering the high number of comparison carried out, there is a significant probability that some tests make errors.

In the case of Ca and Mg (Fig. 4), the robust $\sigma$ of all methods are lower than ca 10% for concentration of between 10 and 100 mg l$^{-1}$. Below 10 mg l$^{-1}$, the EDTA titration shows a distinctly and generally significantly lower precision. At around 0.1-0.3 mg l$^{-1}$, the performances of AAS, IC and ICP are still comparable, but the robust $\sigma$ are higher (19-29% for Ca, 10-26% for Mg).

![Fig. 1. Relative standard deviation of the results of each intercomparison exercise vs the median value, for base cations (a) and ammonium and anions (b).](image-url)
Fig. 2. Percentage of the results considered as outliers for each ion and intercomparison exercise vs the median value, for base cations (a) and ammonium and anions (b).

Fig. 3. Relative standard deviation of the results of each intercomparison exercise, after outlier rejection, vs the median value, for base cations (a) and ammonium and anions (b).
In the case of Ca, only in the first exercises, the few laboratories using IC obtained a robust $\sigma$ at low concentration (56% at 0.21 mg l⁻¹) significantly higher than that obtained using AAS. For Mg levels of below 0.3 mg l⁻¹, in some cases the results obtained using IC or ICP gave robust $\sigma$ slightly, but significantly, higher than using AAS.

All methods used for sodium analysis (Fig. 5a) gave results with robust $\sigma$ lower than 13% for levels higher than 1 mg l⁻¹. However, the precision of AES was lower than that of all other methods used, mainly for samples with concentrations of 0.7 mg l⁻¹ or less. Also in this case, the few laboratories using IC in the first exercise obtained a high robust $\sigma$ (40% and 22% at 0.17 and 1.2 mg l⁻¹, respectively).

Also in the case of potassium (Fig. 5b), AAS, IC and ICP perform better than AES in all exercises. However, for levels lower than 0.25 mg l⁻¹, IC gives results with higher robust $\sigma$. In contrast, AES shows a lower robust $\sigma$, but a higher proportion of outlying results.

The Nessler spectrophotometric method is clearly not suitable for ammonium analysis at a concentration lower than 1.3 mg N l⁻¹ (Fig. 6a). Apart from continuous flow analysis using indophenol blue, the performances of all the other methods used compare well at the lower levels. At above 1 mg N l⁻¹ the indophenol blue method and both continuous flow methods often give robust $\sigma$ significantly lower than those given by IC and ion selective electrode.

All the samples used for the intercomparison exercises had sulphate concentration higher than 1 mg l⁻¹. IC gave results with robust $\sigma$ lower than 8% (Fig. 6b) while the precision of the turbidimetric methods was significantly lower even at the highest concentration levels. Continuous flow analysis (CF) was used only once by more than five laboratories, and the robust $\sigma$ resulted around 15% at 1.7 mg l⁻¹, significantly higher than those obtained using IC.

Of the five methods commonly used for nitrate analysis (Fig. 7a), IC was the most popular and the best performer, with robust $\sigma$ of between 4 and 8% for levels higher than 1 mg N l⁻¹, and lower than 16% even in the most diluted samples. Spectrophotometry after Cd reduction, and UV-spectrophotometry at 220 nm generally give results with significantly higher robust $\sigma$, while continuous flow analysis by Cd reduction and the phenoldisulphonic acid spectrophotometric method are comparable to IC in most, but not all the exercises.
Fig. 4. Robust relative \( \sigma \) of the results obtained using each method for Ca (a) and Mg (b) analysis, in relation to the expected concentration of the sample. Legend as in table 2. Filled symbols denote robust \( \sigma \) values significantly higher than those of the results obtained through the most used method, \( F \)-test, \( p<0.01 \).

Fig. 5. Robust relative \( \sigma \) of the results obtained using each method for Na (a) and K (b) analysis, in relation to the expected concentration of the sample. Legend as in table 2. Filled symbols denote robust \( \sigma \) values significantly higher than those of the results obtained through the most used method, AAS \( F \)-test, \( p<0.01 \).
Fig. 6. Robust relative σ of the results obtained using each method for the analysis of ammonium (a) and sulphate (b), in relation to the expected concentration of the sample. Legend as in table 2. Filled symbols denote robust σ values significantly higher than those of the results obtained through the most used methods, SPEC for ammonium and IC for sulphate (F-test, p<0.01).

Fig. 7. Robust relative σ of the results obtained using each method for the analysis of nitrate (a) and chloride (b), in relation to the expected concentration of the sample. Legend as in table 2. Filled symbols denote robust σ values significantly higher than those of the results obtained through the most used method, IC (F-test, p<0.01).
In the case of chloride (Fig. 7b), all the methods used by more than five laboratories give robust σ ranging between 8 and 12% at 8.3 mg l\(^{-1}\). However, IC, CIA and CF perform significantly better than the other methods for most of the diluted samples. More than 80% of the laboratories in any exercise used IC for chloride analysis, and the robust σ obtained is lower than 10% for concentration higher than 1 mg l\(^{-1}\), and ranges between 13 and 41% for concentrations of between 0.17 and 0.61 mg l\(^{-1}\). At the lowest levels, all methods show low precision, and this may reflect the difficulty of handling such diluted solutions without affecting their chloride content.

4. CONCLUSIONS

The results of the nine intercomparison exercises carried out in the framework of the AQUACON project highlight the unreliability of some analytical techniques in the range of concentrations typically found in rainwater and freshwater. Of the methods used by more than five laboratories, the following proved to be unreliable: EDTA titration (for Ca and Mg), atomic emission spectrometry (for Na and K), ion capillary electrophoresis (for K), the Nessler spectrophotometric method (for ammonium), and continuous flow analysis (for sulphate), spectrophotometry after Cd reduction and UV-spectrophotometry (for nitrate) and HgNO\(_3\) colorimetric titration with diphenilcarbazone, ion selective electrode, Hg(SCN)\(_2\) spectrophotometry with ferric ion and AgNO\(_3\) titration with K\(_2\)CrO\(_4\) as indicator (for chloride).

However, outlying results are also produced by laboratories using regularly well-performing methods. To improve precision, it is therefore important to adopt and routinely apply well-defined procedures of analytical quality control (AQC).

Within the laboratory, AQC procedures must include every aspect of the laboratory facilities, as well as training for the technicians, a handbook containing detailed descriptions of the procedures which is readily accessible to the analyst as s/he works, and a regular check of the equipment used as well as the quality of chemicals and de-ionised water. Other aspects of basic importance in the AQC are calibration procedures and the systematic use of blank charts and control charts, the latter using stabilised solutions with concentration values in the range of the samples generally analysed (Garfield 1991; A.P.H.A. 1995). AQC procedures external to the laboratory include the use of reference material and participation in interlaboratory studies (Quevauviller & Maier 1994).

A further important step in improving analytical quality is that of checking the ion balance of the solution, which is possible when the concentrations of all major ions are measured, as in the case of the samples considered in this paper. Under standard analytical conditions and with a low content of total organic carbon (TOC), the difference between the sums of the total concentration of anions and that of cations (in µeq l\(^{-1}\)) should not exceed 3-5% of the average of the concentration of anions and that of cations.

A further check is a comparison between measured conductivity and the value of conductivity obtained by the sum of the concentration of each ion multiplied by its equivalent conductance. The difference between measured and calculated conductivity should be lower than 5%. A correction for ionic strength is needed for conductivity values above about 15-20 mS m\(^{-1}\). Provided that low pH and high TOC would not interfere, dispersion plots of the sum of cations and the sum of anions vs conductivity are also useful to identify at one glance whether the concentration of a cation or that of an anion is erroneously determined.

Detailed checking procedures are presented in most analysis handbooks (e.g., A.P.H.A. 1995). They are also discussed in the reports of the AQUACON exercises, to promote their adoption in routine procedures.

As an example, if all laboratories would check ion balance and the agreement between measured and calculated conductivity before sending the results, 94% and 98% of the outlying results would be corrected, assuming a threshold of 10% and 5%, respectively.

After outlier rejection or accommodation, at concentration levels higher than 1 mg l\(^{-1}\), it is possible for most of the ions to obtain a standard deviation between laboratories of around 10%, i.e. a reproducibility value of about 25-30% (ISO 1994). However, at lower levels the value of σ and the reproducibility limit rapidly increase (Fig. 2). This pattern underlines that the comparability of results obtained by different laboratories must be considered with care, in particular when interpreting interregional or international studies and/or mapping. Interlaboratory exercises should be an essential part of these studies, with the aim of evaluating and improving the comparability of the data produced.

ACKNOWLEDGMENTS

Dr. Anders Wilander (Institutionen för Miljöanalys, Sveriges Lantbruksuniversitet) and Dr Stefano Polesello (CNR Istituto di Ricerche sulle Acque) provided useful criticism and greatly improved the quality of the manuscript.

REFERENCES


Received: December 1998
Accepted: March 1999


