Shipboard analytical intercomparison of dissolved iron in surface waters along a north-south transect of the Atlantic Ocean
Bowie, AR; Achterberg, EP; Blain, S; Boye, M; Croot, PL; de Baar, HJW; Laan, P; Sarthou, G; Worsfold, PJ
Published in:
Marine Chemistry

DOI:
10.1016/S0304-4203(03)00091-4

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2003

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Download date: 20-09-2020
Shipboard analytical intercomparison of dissolved iron in surface waters along a north–south transect of the Atlantic Ocean

Andrew R. Bowiea,*, Eric P. Achterberga, Stéphane Blainb, Marie Boyec,d, Peter L. Crooef, Hein J.W. de Baarc, Patrick Laanc, Géraldine Sarthoub, Paul J. Worsfolda

aDepartment of Environmental Sciences, Plymouth Environmental Research Centre, University of Plymouth, Plymouth PL4 8AA, UK
bInstitut Universitaire Européen de la Mer, Université de Bretagne Occidentale, Place Nicolas Copernic, 29280 Plouzane, Brest, France
cNetherlands Institute for Sea Research (NIOZ), Postbus 59, 1790 AB Den Burg, Texel, The Netherlands
dDepartment of Marine Biology, University of Groningen, PO Box 14, 9750 AA Haren, The Netherlands

Received 21 May 2002; received in revised form 5 May 2003; accepted 5 May 2003

Abstract

A shipboard analytical intercomparison of dissolved (<0.2 μm) iron in the surface waters of the Atlantic Ocean was undertaken during October 2000. A single underway surface (1–2 m) seawater sampling and filtration protocol was used, in order to minimise differences from possible sample contamination. Over 200 samples (1/h) were collected over 12 days and analysed immediately using four different analytical methods, based on three variants of flow injection with luminol chemiluminescence (FI–CL) and cathodic stripping voltammetry (CSV). Dissolved iron concentrations varied between 0.02 and 1.61 nM during the intercomparison. On average, CSV [Electroanalysis 12 (2000) 565] measured 0.08 nM higher iron concentrations than one FI–CL method [Anal. Chim. Acta 361 (1998) 189], which measured 0.13 nM higher iron values than the other two [Anal. Chem. 65 (1993) 1524; Anal. Chim. Acta 377 (1998) 113]. Statistical analyses (paired two-tailed t-test) showed that each analytical method gave significantly different dissolved iron concentrations at the 95% confidence interval. These data however, represent a significant improvement over earlier intercomparison exercises for iron. The data have been evaluated with respect to accuracy and overall inter-laboratory replicate precision, which was generally better than the 95% confidence intervals reported for the NASS Certified Reference Materials. Systematic differences between analytical methods were probably due to the extraction of different physico-chemical forms of iron during preconcentration, either on the microcolumn resin (in the FI methods) or with competing ligand equilibration (in the CSV method). Small systematic concentration differences may also have resulted from protocols used for quantification of the analytical blank and instrument calibration.

© 2003 Elsevier B.V. All rights reserved.

Keywords: Dissolved iron; Seawater; Analytical intercomparison; Atlantic Ocean

1. Introduction

The importance of iron for primary production in the ocean (Martin and Fitzwater, 1988; Boyd et al., 2000) and global carbon cycling (Watson et al., 2000)
requires that it is measured routinely and accurately as part of any attempt to understand the factors controlling the functioning of marine ecosystems. This need has led to the rapid development of several shipboard and land-based analytical methods capable of measuring dissolved iron at sub-nanomolar levels in seawater. At the 1998 international symposium of SCOR-IUPAC Working Group 109 on the Biogeochemistry of Iron in Seawater, it was reported that concentrations of iron in surface seawater throughout the world varied over several orders of magnitude. The lack of rigorous intercomparison exercises and appropriate seawater reference materials (NASS-4 is $1.88 \pm 0.29$ nM, NASS-5 is $3.71 \pm 0.63$ nM and BCR CRM 403 is not certified for iron) means that the scientific community has little ability to correlate these observations or to distinguish between environmental variability and analytical data quality. It is widely recognised that a tremendous amount has been learned about the distribution and biogeochemistry of iron in recent years, but the ability to relate one study to another remains problematic.

There have been very few attempts at analytical intercomparison for trace metals to date. The intercalibration exercise organised by the International Council for the Exploration of the Sea (ICES; Bewers et al., 1981) was largely unsuccessful for iron, with a high inter-laboratory precision reported for analysis of acidified samples in the range $14.5–31.5$ nM, concentrations two orders of magnitude greater than those present in surface seawater. During the first Intergovernmental Oceanographic Commission (IOC) baseline survey in April 1990 (Landing et al., 1995), only three laboratories reported data for iron during an intercomparison for a station in the eastern Atlantic Ocean, with a 2- or 3-factor degree of variability between concentrations (ranging from 0.1 to 5.3 nM). Iron measurements were performed during the 1996 IOC baseline cruise from Uruguay to Barbados (Vink and Measures, 2001; Powell and Donat, 2001), although no intercomparison data for iron has been published. Unfortunately, time and space constraints on-board ships often mean that analytical method intercomparison exercises are rarely performed at sea. There is thus an urgent need for standardisation of sampling and analytical methods in order to ensure the highest possible integrity, reliability and comparability of reported iron data. Clearly, with so many sampling and analytical variants in current use, it is difficult to attribute any differences in reported concentrations to one particular step in the overall process without first determining differences between analytical methods.

In October 2000, four groups, using different analytical methods, monitored dissolved iron concentrations in surface seawater during a north–south transect of the eastern Atlantic Ocean covering approximately $50^\circ$ of latitude ($27^\circ$N to $19^\circ$S). This paper presents the results from the shipboard intercomparison of the investigators’ analytical methods during this high resolution (every 1 h) monitoring of dissolved iron on fresh samples collected during the cruise, the first multi-investigator exercise to take place at sea. To minimise differences in concentrations that may result from low-level contamination during sample collection, one standard underway sampling protocol was used throughout, based on a towed fish connected to a trace metal clean pumping system. Individual laboratories were responsible for subsample preservation (e.g. acidification), pretreatment (e.g. reduction, pho-

![Fig. 1. Cruise track taken during a north–south transect of Polarstern (Anreise expedition, ANT XVIII/1). Samples for the iron intercomparison exercise were collected between a port-call at Las Palmas (26.95°N, 16.06°W) and 19.09°S, 5.07°E.](image-url)
to-oxidation) and analysis, which was based on flow injection–chemiluminescence (FI–CL) and competitive ligand equilibration–cathodic stripping voltammetry (CLE–CSV) methods. The principle aim of this study was to compare the accuracy and precision of the four shipboard methods, to determine whether comparable results were significantly different at the 95% confidence interval and to consider the reasons for any systematic bias between methods.

2. Materials and methods

2.1. Sampling

Sampling and analyses were undertaken during voyage ANT XVIII/1 (September 29th to October 23rd 2000) on-board R/V Polarstern, on a north–south transect of the eastern Atlantic Ocean from Bremerhaven (Germany) to Cape Town (South Africa) (cruise Anreise; Fig. 1). Surface water samples were collected every hour between 19:00 (UTC) on 8th October and 09:00 on 19th October 2000 (26.95°N, 16.06°W to 19.09°S, 5.07°E), apart from six short periods during the transect when the sampling unit was recovered from the water to repair a partially collapsed inlet tube.

Underway sampling of surface (1–2 m) seawater was performed using a towed polyurethane-coated torpedo-shaped fish (1 m long, 50 kg weight), fitted with a Teflon FEP nose tube and deployed off the crane arm of a hydrographic winch at distance of ~ 5 m from the ship’s starboard side (de Jong et al., 1998; Bowie et al., 2001). The fish was capable of being towed up to speeds of 14.5 knots. Seawater was pumped on-board through acid-washed braided PVC tubing using a variable speed high volume peristaltic pump (model 7591-00, Cole Palmer Instrument, Hanwell, UK), fitted with silicone pump tubing and filtered through a Sartobran-P polypropylene cartridge unit with 0.2 μm cellulose acetate filter membrane (Sartorius, Epsom, UK). Water from the sampling tubing passed through a flow regulator and entered a sink in a class-1000 clean container laboratory positioned on the ship’s aft deck.

<table>
<thead>
<tr>
<th>Group code</th>
<th>UBO</th>
<th>NIOZ</th>
<th>UoP</th>
<th>CSV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Affiliation*</td>
<td>UBO</td>
<td>NIOZ</td>
<td>UoP</td>
<td>HDPE, 100 ml (Nalgene)</td>
</tr>
<tr>
<td>Subsample bottle type</td>
<td>LDPE, 60 ml (Nalgene)</td>
<td>LDPE, 100 ml (Emergo)</td>
<td>NIOZ and UoG</td>
<td></td>
</tr>
<tr>
<td>Washing protocol</td>
<td>Decin bath, 5% (1 week)</td>
<td>Decin bath, 5% (1 week)</td>
<td>Decon bath, 5%, hot (1 week)</td>
<td></td>
</tr>
<tr>
<td>Step 1</td>
<td>Decin bath, 5% (1 week)</td>
<td>Decin bath, 5% (1 week)</td>
<td>Decon bath, 5%, hot (1 week)</td>
<td></td>
</tr>
<tr>
<td>Step 2</td>
<td>HCl fill, pro analysis, 6 M (1 week), outside wall rinsed with MQ water</td>
<td>HCl fill, Analytical grade, 6 M, immersed in hot (60 °C) MQ water bath (3 days)</td>
<td>HCl bath, Aristar grade, 6 M (2 weeks)</td>
<td>HNO3 bath, Aristar grade, 3 M (2 weeks)</td>
</tr>
<tr>
<td>Step 3</td>
<td>HCl fill, Suprapur, 6 M (3 days)</td>
<td>–</td>
<td>HNO3 bath, Aristar grade, 3 M (2 weeks)</td>
<td>HNO3 bath, AnalaR grade, 2 M (1 week)</td>
</tr>
<tr>
<td>Storage</td>
<td>HCl fill, Suprapur grade, 0.01 M, triple bagged</td>
<td>Q-HCl fill, 0.1 M, double bagged</td>
<td>Q-HCl fill, 0.01 M, double bagged</td>
<td>Q-HCl fill, 0.01 M, double bagged</td>
</tr>
</tbody>
</table>

All sample bottles were thoroughly rinsed (3 ×) with copious amount of MQ water in a Class-100 clean laboratory between each washing stage.

*UBO: Universite de Bretagne Occidentale (France); NIOZ: Royal Netherlands Institute for Sea Research (The Netherlands); UoP: University of Plymouth (UK); UoG: University of Groningen (The Netherlands).
2.2. Sample collection

All sampling bottles were thoroughly washed prior to use and rinsed with copious amounts of Milli-Q (MQ) water, with each group following their standard procedures (Table 1). For collection, a 1-l wide-mouth PTFE bottle (Nalgene) was rinsed three times with filtered seawater from the underway supply, filled, closed and gently shaken. This bottle was immediately transferred into a class-100 laminar flow hood and subsampled into four smaller LDPE or HDPE bottles (variable volumes), which were provided by each participant (Table 1). Each subsample was placed in a double zip-locked bag and stored at room temperature (<1 h, FI–CL methods) or in the fridge (≤4 °C; up to 48 h, CSV method) prior to pretreatment (e.g. acidification, UV oxidation, reduction) and analysis by each investigator. Clean room garments (overalls, hats and boots) and polyethylene gloves were worn at all times by personnel handling the sample collection bottle, the filtering equipment and the subsampling bottles. All possible precautions were taken to prevent contamination during sampling and analysis.

2.3. Analysis

The techniques and procedures used for iron analyses by each group are summarised in Table 2. Instrument calibration for the FI methods was achieved by

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Analytical methods used by the research groups</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Group</strong></td>
<td><strong>UBO</strong></td>
</tr>
<tr>
<td><strong>Sample acidification</strong></td>
<td>Q-HCl, 0.01 M, pH ~ 2</td>
</tr>
<tr>
<td><strong>Sample pretreatment</strong></td>
<td>Natural oxidation in dark, 1 h</td>
</tr>
<tr>
<td><strong>Storage</strong></td>
<td>Room temperature, bagged</td>
</tr>
<tr>
<td><strong>Analysis</strong></td>
<td>FI–CL, FeIII, luminol, H₂O₂</td>
</tr>
<tr>
<td><strong>Preconcentration</strong></td>
<td>8HQ immobilised on Toyopearl TSK HW75F (fine) resin (pH 4.5)</td>
</tr>
<tr>
<td><strong>Column size</strong></td>
<td>50 × 4 mm PTFE tubing</td>
</tr>
<tr>
<td><strong>Calibration</strong></td>
<td>Standard curve, 0.5–2.0 nM (n = 4) P (Spectrosol)</td>
</tr>
<tr>
<td><strong>Iron standard</strong></td>
<td>1000 mg l⁻¹ FeIII atomic absorption standard (Spectrosol)</td>
</tr>
<tr>
<td><strong>Time between sampling and pretreatment</strong></td>
<td>&lt;1 h</td>
</tr>
<tr>
<td><strong>Time between pretreatment and analysis</strong></td>
<td>1 h</td>
</tr>
<tr>
<td><strong>Time for one analytical cycle</strong></td>
<td>5 min</td>
</tr>
<tr>
<td><strong>Reference</strong></td>
<td>Obata et al., 1993</td>
</tr>
</tbody>
</table>
obtaining a standard curve over the range 0.5–2.0 nM using acidified low iron seawater collected along the transect. Each group used their own batch of low iron seawater, which contained <0.2 nM Fe. A new standard curve was obtained for each batch of reagents or whenever there was a significant change in system sensitivity (e.g. with laboratory temperature changes). The CSV method was calibrated using standard additions to each sample over the range 0.5–2.0 nM. The UBO, UoP and CSV groups measured peak heights whereas the NIOZ group determined peak areas. The NIOZ, UBO and CSV groups prepared their iron standards independently by dilution of a 1000 mg l\(^{-1}\) iron(III) atomic absorption standard (Spectrosol, Merck). The UoP FI–CL method is based on sample reduction of iron(III) to iron(II) and standards were therefore prepared by serial dilution of a 0.02 M \((\text{NH}_4)_2\text{Fe} (\text{SO}_4)_2\cdot 6\text{H}_2\text{O}\) solution in 0.1 M Q-HCl. Replicate analyses \((n = 3\) or \(n = 4\)) were performed for each sample and standard solution. Suspect or contaminated samples (based upon: (1) significant differences between samples run immediately before and after the suspect sample, (2) poor precision between replicate peaks, or (3) a lack of oceanographic consistency with surface hydrography) were re-analysed. Contaminated samples were then rejected at the individual discretion of each group.

During the cruise, each FI–CL group closely followed their published methods (Table 2). The only minor modifications were that the UBO group purified their luminol reagent by passing it through an 8-hydroxyquinoline (8HQ) column and the UoP group added Na\(_2\)SO\(_3\) reducing agent (100 \(\mu\)M final concentration) to their 400 nM iron(II) working standard. The CSV method was an adaptation of the one reported by Croot and Johansson (2000), using the synthetic ligand 2-(2-thiazolylazo)-p-cresol (TAC). Here, 20 ml of the sample was UV digested (using a 600-W high-pressure mercury vapour lamp) at ambient seawater pH for 4 h. No oxidant was added to the sample in order to minimise the analytical blank. Voltammetric measurement was carried out at pH 8.05 using borate buffer (final concentration 5 mM) with 10 \(\mu\)M TAC. A deposition potential of \(-0.4\) V was applied for 150 s, during which time the solution was stirred to facilitate the adsorption of the Fe–TAC\(_2\) complex onto the Hg drop. The voltammetric procedure was carried out using the fast linear sweep waveform from \(-0.4\) to \(-0.9\) V at 10.1 V s\(^{-1}\) (step potential 1.98 mV) and the stripping reduction current measured. Each scan was repeated three times. Due to the extended analysis time compared to the more rapid FI methods, a sample for CSV determination was taken from the underway supply approximately every 4 h.

### 3. Results and discussion

#### 3.1. Analytical performance

During the transect, a total of 219 surface samples were collected over a 12-day period. This large sample set enabled us to observe systematic errors in the presence of random errors for individual measurements. The biogeochemical interpretation of these data will be reported elsewhere (Sarthou et al., submitted for publication). The analytical figures of merit for the methods used are given in Table 3. Procedural blanks, detection limit (DL) and precision of each method are depicted in Fig. 2. The definition of the blank varied for each investigator. This practice was deliberately adopted to be consistent with the historical methods and previously reported oceanographic data. The blank for the UBO group was obtained from the signal given during a 5-s loading of MQ water on the 8HQ column (reagent blanks were lower than the DL and deemed negligible). The blank signal for the UoP group was defined as the signal given during a 1 min loading of sample buffer only (NH\(_4\)OAc) on the 8HQ column (followed by a routine 40 s MQ water rinse), plus that given by the addition of 100 \(\mu\)M Na\(_2\)SO\(_3\) and 0.01 M Q-HCl. The latter was measured by double spiking a low-iron concentration sample with extra acid and reducing agent. The NIOZ group blank was determined by measuring the signal given by loading the 8HQ column with MQ water for 1 min, plus the signal given by the reagents (acid and buffer; confirmed by double spiking a low iron concentration sample). The blank signal for the CSV group was determined by linear regression after the addition of excess borate buffer (at two-fold and four-fold the working 5 mM addition) to low-iron UV-digested filtered seawater, and by the determination of iron in the TAC reagent (by GFAAS). The DL for all groups was defined as three times the standard deviation for
Replicate analyses \((n = 4)\) of the blank. A summary of all iron data collected by each group during the intercomparison exercise is given in Table 4. A small number of samples were “lost”, as the result of insufficient sample water and/or instrument malfunction. Contaminated outlier samples were excluded from data used for statistical calculations and interpretation. Outliers were rejected at the discretion of the research laboratories, which was generally based upon an inconsistent oceanographic trend in the surface iron distribution (see Section 2.3).

### 3.2. Accuracy checks

The accuracy of the FI–CL methods was previously ascertained in home laboratories and on earlier cruises by analysing the seawater CRMs NASS-4 and/or NASS-5, the best option currently available. How-

![Fig. 2](image-url)  
**Fig. 2.** Procedural blanks (pM), detection limits (pM) and analytical precision (RSD, %) of the methods employed during the intercomparison. Error bounds indicate ± one standard deviation of all measurements made during the transect (no error bounds are quoted for the detection limit of the CSV method).
ever, it should be emphasized that the iron concentrations certified in these materials are up to two orders of magnitude greater than open-ocean waters (and thus not representative of the concentrations found in this study). Results (Table 5) show reasonable agreement with certified values and t-tests show there is no evidence of systematic error (95% confidence interval). Since the CSV measurements are made in seawater at neutral to basic pH, and reference seawater is acidified to pH 1.6 using Q-HNO3, it was inappropriate (and unrepresentative of field measurements) to analyse a NASS seawater standard using this method.

3.3. Instrument reliability

Each system was assembled for shipboard use and initial calibrations completed within 24 h of departure from port. Minor problems (e.g. partially blocked flow lines) were quickly rectified and resulted in only short down-times. One batch of reagents typically lasted 9 and 12 h for the UoP/NIOZ and UBO FI–CL methods, respectively. Two temperature-controlled (20 °C) clean container laboratories were used to minimise changes in system sensitivity. However, in the Equatorial region, a problem with the air-conditioning system in one container housing the UBO and NIOZ analysis systems resulted in a temporary temperature increase to 37 °C. This resulted in 73% and 45% rises in the sensitivities of the UBO (from 59.4 to 102.8 × 10^5 counts nM^(-1)) and NIOZ (from 84.5 to 122.4 × 10^5 counts nM^(-1)) methods, respectively. UBO and NIOZ methods were calibrated regularly during this period (samples #78–105) to compensate for the sensitivity change. In addition, a problem with a contaminated injection value on the UBO FI–CL system resulted in absence of their data for 17 consecutive samples (#163–179).

3.4. General observations

The surface distribution of dissolved iron along the transect (all data) is shown in Fig. 3a. Values ranged from 0.02 nM (UBO) in the South Atlantic up to 1.61 nM (UoP) to the west of Africa. The range of dissolved iron concentrations measured during this intercomparison was generally consistent with the eastern Atlantic data of Vink and Measures (2001) (0.4–1.4 nM), but noticeably lower than the unfiltered data of Powell et al. (1995) and Bowie et al. (2002). A similar trend in concentrations was observed by each group, despite methodological differences. Fig. 3b shows the mean ± 1 standard deviation for all reported values. At the start of the transect, concentrations generally decreased from >1.0 to ~0.3 nM between samples #030 and 050, then increased to ~0.6 to 0.8 nM between samples #079 and #105 before tending towards a baseline concentration of ~0.1 to 0.3 nM during the latter half of the exercise. A smaller subset (every 10th sample) of these data is shown in Fig. 3c to highlight the relative differences between methodologies. In general, reported iron concentrations increased in the order NIOZ ≈ UBO < UoP < CSV (Table 4), although consistent offsets were not ob-

### Table 4

Summary of surface dissolved iron (nM) data collected by each group during the shipboard intercomparison

<table>
<thead>
<tr>
<th>Group</th>
<th>UBO</th>
<th>NIOZ</th>
<th>UoP</th>
<th>CSV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum</td>
<td>0.02</td>
<td>0.03</td>
<td>0.05</td>
<td>0.20</td>
</tr>
<tr>
<td>Maximum</td>
<td>1.11</td>
<td>1.11</td>
<td>1.61</td>
<td>0.97</td>
</tr>
<tr>
<td>Average</td>
<td>0.278</td>
<td>0.275</td>
<td>0.414</td>
<td>0.490</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>0.207</td>
<td>0.246</td>
<td>0.286</td>
<td>0.218</td>
</tr>
<tr>
<td>No. of measurements</td>
<td>194</td>
<td>211</td>
<td>212</td>
<td>52</td>
</tr>
<tr>
<td>No. of samples “lost”</td>
<td>20</td>
<td>6</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>No. of samples contaminated</td>
<td>5</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

Samples for FI methods were taken approximately every 1 h, whereas samples for CSV were taken approximately every 4 h.

### Table 5

Results of the analysis of iron (in nM) in North Atlantic open-ocean seawater reference materials

<table>
<thead>
<tr>
<th>Standard solution</th>
<th>Certified value</th>
<th>UBO</th>
<th>NIOZ</th>
<th>UoP</th>
<th>CSV</th>
</tr>
</thead>
<tbody>
<tr>
<td>NASS-4</td>
<td>1.88 ± 0.29</td>
<td>2.05 ± 0.11</td>
<td>1.90 ± 0.21</td>
<td>2.02 ± 0.17</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>(n = 3)</td>
<td>(n = 7)</td>
<td>(n = 9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NASS-5</td>
<td>3.71 ± 0.63</td>
<td>3.52 ± 0.07</td>
<td>3.35 ± 0.51</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>(n = 5)</td>
<td>(n = 8)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Certified values are based on three independent methods of analysis. Uncertainties represent 95% confidence limits for an individual subsample. No data is available for the CSV method.

*a Mean of means, first reported in de Jong et al. (2000).

b N/A = not analysed.
vious for every sample. As expected, methodological difficulties in accurately measuring exceptionally low concentrations of iron resulted in a greater discrepancy between methods at relatively low compared with higher dissolved iron concentrations, with a RSD of the mean of all methods of 26% for samples containing >0.6 nM and 74% for samples containing <0.2 nM iron.

Six regression graphs for the direct two-way comparison between each dataset are shown in
The product-moment correlation coefficient ($r^2$) was consistent between each set of comparisons and varied between 0.59 for UoP vs. NIOZ and 0.74 for CSV vs. UBO. The positive intercept of the regression line in Fig. 4b–f approximately equates to the difference (in concentration units) of the UoP and CSV methods relative to the UBO and NIOZ methods. There was good agreement between the NIOZ and UBO datasets ($\text{NIOZ} = \text{1.12} \times \text{UBO} - \text{0.01}$, $r^2 = 0.70$, $n = 188$; Fig. 4a).

3.5. Statistical analysis

In order to test whether there was a significant difference between results obtained by each analytical method, paired two-tailed $t$-tests were performed. The concentration of iron in each sample varied in the range 0.02–1.61 nM. Here, we assume that errors, either random or systematic, are independent of concentration. Interestingly, results (Table 6) showed that each analytical method gave significantly different iron concentrations at the 95% confidence interval ($P = 0.05$).

Of the 219 samples collected during the exercise, only 43 were analysed by every group. In order to remove any bias in the following calculations, only data from this smaller subset of samples is used. Since the subset represents approximately every 5th sample collected along the transect, it is assumed to be representative of the complete dataset.

Fig. 5 shows the absolute difference of each group’s measurements from the mean for each sample in this subset. The mean and standard deviation between the four investigators for each of the 43 samples was calculated and the minimum, average and maximum of these values used to determine the overall inter-laboratory replicate precision for each sample in this subset. This statistic, expressed as two times the standard deviation ($2 \times \text{S.D.}$) of the mean of the reported values, was expressed in concentration units (Table 7). The overall replicate precision ($2 \times \text{S.D.}$) among investigators ranged from 0.15 to 0.48 nM and was generally better than the 95% confidence intervals reported for the CRMs, which were 0.29 and 0.63 nM for NASS-4 and NASS-5, respectively. The overall precision reported here for
the intercomparison exercise represents the intra- and inter-laboratory variance due to subsampling, subsample bottle preparation, sample preservation and analysis, whereas the reported precision of the CRMs is based upon analytical variance only.

The relative accuracy of the results from each group was evaluated by calculating the root-mean-square (RMS) deviations of the reported values from the mean values. This was done to allow a direct comparison with the statistical parameters used in the previous intercomparison for trace metals (Landing et al., 1995). For example, for the UBO group, this would be:

\[
RMS = \sqrt{\frac{\sum_{i=1}^{n} (x_{UBO} - \bar{x}_{All})^2}{n}}
\]

where \(x_{UBO}\) is the UBO data for sample \(i\), \(\bar{x}_{All}\) is the mean of the data for all four groups for sample \(i\) and \(n\) is the subset of 43 samples. This statistic was calculated for each group and multiplied by \(2 \times \text{RMS}\) to be consistent with the approach used by Landing et al. (1995). These data are also equal to \(2 \times \text{S.D.}\) for a large sample set (e.g. \(n > 11\)). The \(2 \times \text{RMS}\) deviation values ranged from 0.23 (UoP) to 0.32 nM (CSV) (Table 7).

### 3.6. Examination of analytical methods

Our intercomparison results demonstrate that over a large population (approximately 200 samples) there

---

**Table 6**

<table>
<thead>
<tr>
<th>Groups</th>
<th>(n)</th>
<th>(t) (critical)</th>
<th>(t) (experimental)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NIOZ/UBO</td>
<td>188</td>
<td>1.97</td>
<td>2.53</td>
</tr>
<tr>
<td>CSV/UoP</td>
<td>50</td>
<td>2.01</td>
<td>4.00</td>
</tr>
<tr>
<td>CSV/NIOZ</td>
<td>50</td>
<td>2.01</td>
<td>9.49</td>
</tr>
<tr>
<td>UoP/NIOZ</td>
<td>206</td>
<td>1.97</td>
<td>9.89</td>
</tr>
<tr>
<td>UoP/UBO</td>
<td>190</td>
<td>1.97</td>
<td>12.95</td>
</tr>
<tr>
<td>CSV/UBO</td>
<td>46</td>
<td>2.01</td>
<td>14.04</td>
</tr>
</tbody>
</table>

---

**Fig. 4.** Six regression plots comparing each of two analytical methods. 1:1 lines are shown (dashed).
was a systematic bias between all methods at the 95% confidence interval. Mean values for the whole transect were $0.28 \pm 0.21 \ (n=194)$, $0.28 \pm 0.25 \ (n=211)$, $0.41 \pm 0.29 \ (n=212)$ and $0.49 \pm 0.22 \ (n=52)$ for UBO, NIOZ, UoP and CSV methods, respectively. The similarity between UBO and NIOZ mean values

Fig. 5. Absolute differences of each analytical method from the mean. Data for 43 samples that were analysed by all four investigators are shown.
is reflected in the observation that these methods were not significantly different at the 99% confidence interval \((P = 0.01)\). Interestingly, previous results from a less rigorous shipboard comparison between FI–CL (NIOZ) and CSV methods (de Jong et al., 2000) showed generally good agreement for 17 samples collected through a vertical profile in the eastern North Atlantic, although there was some scatter in the data. Dissolved iron concentrations were in the range 0.5–1.6 nM.

During our study, samples were collected using one standard protocol from an underway sampling system but random differences between the iron concentrations reported by each group may have been introduced for individual samples through low-level contamination during the subsampling process. Furthermore, discrepancies between investigators may have resulted from protocols used for quantification and subtraction of the analytical blank. This ranged between 0.02 and 0.17 nM (Table 3), and thus represented an important fraction (18–54%) of the analytical signal during the second half of the transect, where surface dissolved iron levels ranged from 0.11 (UBO) to 0.32 nM (CSV) (mean of data for each group for samples #106–219). It is unlikely however, that the higher iron concentrations obtained using the CSV method can be explained by blank subtraction, since the CSV blank was also the highest (Table 3). In the future, the iron community may wish to recommend a “best practice” definition for the analytical blank.

In addition, discrepancies between methods may have resulted from variations in system sensitivities. The RSD of the calibration slope (for each method) during the intercomparison varied from 14.1% (UoP) to 35.9% (CSV) (Table 3). Each FI method was calibrated by obtaining a standard curve, whereas the CSV method was calibrated using standard additions to each sample. Instruments were re-calibrated frequently (at least once per day or for each batch of reagents), resulting in typically one calibration for 10–20 samples. There may, however, have been small changes in sensitivity during operation due to temperature fluctuations (affecting both the PMT detector and CL chemistry), reagent ageing, degradation of pump tubing quality (reducing flow rates) or subtle matrix effects (affecting loading of iron onto the 8HQ column). Only individual standard additions can ensure a calibration slope is achieved for each sample, but such methods are time-consuming and require a significantly greater volume of sample.

Despite the generic nature of the instrumentation used for the FI–CL methods, there are distinct variations in reaction chemistries between the UBO, NIOZ and UoP methods. The UoP method is based upon reduction of iron(III) to iron(II) in the sample using an excess of Na2SO3, followed by the iron(II)-catalysis of luminol using dissolved oxygen as the oxidant (Bowie et al., 1998; developed from the batch method of Seitz and Hercules, 1972). The UBO and NIOZ methods are based upon the natural oxidation of iron(II) to iron(III) in the sample in the dark (1 h), followed by the iron(III)-catalysis of luminol in the presence of added H2O2 (Obata et al., 1993; de Jong et al., 1998). Both methods rely on the extraction and preconcentration of the analyte from the seawater matrix using 8HQ immobilised on a chemically resistant vinyl polymer resin (Toyopearl TSK) (Landing et al., 1986; Dierssen et al., 2001). Previous work has shown that iron(III) is quantitatively extracted from seawater by 8HQ at pH>3 (de
Jong et al., 1998), whereas iron(II) is quantitatively extracted between pH 4.5 and 8 (Obata et al., 1993; Bowie et al., 1998). Different flow rates and buffering conditions (NH₄OAc; initial pH 5.0–5.5) are thus used for the two methods to obtain the appropriate final loading pH.

8HQ is known to complex strongly with iron with stability constants (log \( K_I \)) for the equilibrium [ML]/[M][L] of 8.71 (0.3 M NaClO₄, 50% dioxane) for iron(II) and 13.69 (0.1 M NaClO₄) for iron(III) (Sillen, 1964; Smith and Martell, 1989). However, in a dynamic flow-through system, equilibrium is unlikely to be reached within the micro-environment of the column and parameters such as sample pH, loading flow rate, eluent concentration, preconditioning, column size and the (in)organic speciation of iron in the sample will affect its complexation on the chelating resin. Importantly, different investigators use 8HQ immobilised on Toyopearl resins of different particle size, porosity and texture (Table 2), which may impact on the efficiency of the extraction, since Toyopearl itself is known to preconcentrate colloidal matter (Landing et al., 1986).

Iron is >99% organically complexed in the South and Equatorial Atlantic Ocean, with ligand concentrations ranging from 0.6 to 2.5 nM (Powell and Donat, 2001). Conditional stability constants with respect to inorganic iron(III) species (log \( K_{FeL,Fe(III)} \)) throughout the ocean range between 19 and 23 (Boye et al., 2001), based on \( \alpha_{Fe} = [Fe^{III}]/[Fe^{II}] = 10^{19} \) (Hudson et al., 1992). The organic side reaction coefficient of iron (\( \mu_{FeL} \)) is equal to \( K_{FeL,Fe(III)} \times \) [ligand] (assuming a Fe/L stoichiometry of 1:1) and thus log \( \alpha_{Fe^{III}-L} \) ranges between 9.8 and 14.4 (log values) for the South Atlantic Ocean. Previous breakthrough capacity experiments (at 95% efficiency) showed that a typical 8HQ column can quantitatively retain up to 56 nmol of iron(II) (Bowie et al., 1998) and 96 nmol of iron(III) (de Jong et al., 1998) in a seawater matrix, compared to the \( \sim 1–10 \) pmol of iron which will typically be loaded onto an 8HQ column during the analysis of an open-ocean sample. Hence, the 8HQ binding sites on a column will not become saturated with dissolved iron (although the 8HQ ligand also has some affinity for seawater matrix ions). A typical 8HQ column is 40 mm long, 3 mm internal diameter (de Jong et al., 2000; Table 2), representing an internal volume of 283 \( \mu \)l. Hence, the concentration of iron(III)-binding sites will be approximately 0.3 mM, resulting in an organic side reaction coefficient for the retention of iron(III) on the 8HQ column (log \( \alpha_{Fe^{III}-8HQ} \)) of 10.2 (assuming a Fe/8HQ stoichiometry of 1:1).

Since the \( \alpha \)-coefficients for Fe(III)-L and Fe(III)-8HQ are of the same order, there will be significant competition between natural organic complexing ligands and the binding sites of the 8HQ column, and the recovery of the iron(III) analysed by FI–CL will be lowered. However, this assumes that the contact time between solution-phase natural organic iron complexes and the solid-phase chelation onto the 8HQ column is long enough to reach the log-phase of the dissociation kinetics of the iron–ligand complex. This is unlikely in rapid flow-through methods. Furthermore, the pH at which the sample is loaded onto the 8HQ column is an important factor affecting the rate of dissociation (natural Fe–L complexes in the sample) and formation (Fe-8HQ on the column) of organic complexes, due to competition for iron by hydrolysis reactions. Despite the widespread use of 8HQ columns for trace metal analyses, there remains uncertainty as to which (in)organic fractions of iron are extracted from seawater and rendered available to the downstream reaction chemistry of the FI method. Further study is required in this regard.

At present, there is no direct analytical method to ascertain the presence and strength of iron(II) chelators in seawater, although indirect evidence based on oxidation rate measurements suggests their existence (Santana-Casiano et al., 2000; Croot et al., 2001). Reference material data for the Bowie et al. (1998) FI–CL method (Table 5) indicates that sample acidification coupled with the addition of excess reducing agent will render iron(II) labile for complexation onto the microcolumn and result in a near-total recovery. Conversely, preliminary experiments suggest that strongly bound iron(III)–organic complexes present in seawater result in a <100% extraction efficiency of iron(III) on 8HQ (Obata et al., 1997; Croot, unpublished data), as predicted from stability constant and \( \mu \)-coefficient data. It is possible that this missing fraction is recoverable as free iron(II) after the addition of excess reducing agent, and thus be measured in the UoP method. For CSV methods, the recovery of organically bound iron will be dependent on the concentration and strength of the complexation of the added ligand with iron (e.g. for TAC, log
\[ K_{\text{FeTAC}}' = 12.4, \text{ Croot and Johansson, 2000} \] and the equilibration time. However, experiments have shown that a 4-h UV digestion step, which precedes the CSV analysis of total dissolved iron, effectively breaks down organically bound iron complexes (Rue and Bruland, 1997; Boye and van den Berg, 2000) and results in its availability to complexation by TAC for the dissolved iron CSV measurement.

### 3.7. Differences due to environmental factors

Atmospheric iron is predominantly associated with deposition of aluminosilicate mineral soil material 1–100 µm in diameter (Jickells and Spokes, 2001). Our measurements were made on filtered (0.2 µm) seawater samples and hence iron bound or adsorbed to dust particles present in seawater was unlikely to be directly determined along the transect. It is evident, however, from this (Sarthou et al., submitted for publication) and other expeditions (Powell et al., 1995; Vink and Measures, 2001; Bowie et al., 2002) that aerosol iron (associated with Al) from the northwest African continent (mainly the Sahara desert) enters the eastern Atlantic Ocean and increases surface iron concentrations. This enrichment in dissolved iron may well be the result of low pH cloud cycling increasing the lability of particulate iron prior to its deposition as heavy rain events in the Inter-tropical Convergence Zone (Zhuang et al., 1990). Soluble iron deposited atmospherically may undergo hydrolysis at the sea surface, resulting in the formation of colloidal material small enough to pass through a < 0.2-µm filter (Nishioka et al., 2001) and bound to colloidal organic ligands present in seawater (Wu et al., 2001). Hence, the intrusive nature and time of sample pretreatment (e.g. acidification, reduction, UV digestion) protocols will render different fractions of colloidal iron present in seawater sample kinetically available to each analytical method. Interestingly, earlier work has shown aged iron hydroxides to be available (88% recovery) to the NIOZ and UBO methods at pH 2, whereas biogenic and sedimentary particles were only partially recoverable (26–28%), even at a pH < 1.5 (Obata et al., 1997).

In addition, although Jickells and Spokes (2001) report an overall solubility of atmospheric iron in seawater at pH 8 to be only 0.8–2.1% of the total iron deposited, a large fraction of this may well occur as bioavailable iron(II), derived from photochemical processes (Zhu et al., 1993). Iron(II) species may be more available to an analytical method based upon the direct measurement of iron(II) (after the addition of excess reducing agent) using luminol CL detection compared with one based upon the natural oxidation of iron(II) to iron(III), especially in the presence of iron(II)-binding ligands (possibly supplied via wet deposition) which may retard natural oxidation rates (Croot et al., 2001). Future work must therefore ascertain the lability of colloidal and organically bound iron to 8HQ extraction since this is a key stage in most shipboard FI systems.

### 4. Conclusions

The surface water distribution of dissolved iron obtained during the intercomparison exercise was in general agreement with previously reported data for this section of the Atlantic Ocean, although this is the first expedition where exceptionally low concentrations (< 0.1 nM) have been observed in the South Atlantic. Due to scatter in the profiles, consistent methodological differences were not always obvious although, on average, reported iron concentrations increased in the order NIOZ ≈ UBO < UoP < CSV. Each analytical method gave significantly different dissolved iron concentrations at the 95% confidence interval (paired two-tailed \( t \)-test).

Systematic discrepancies between methods were due to either (in decreasing level of importance): (1) efficiency of the extraction of iron from the seawater matrix during precollection (resulting in different methods measuring different fractions of iron), (2) errors in the quantification of the analytical blank, and (3) inaccuracies in system calibration. Random differences for individual samples were thought to be due to low-level contamination during subsample processing. The different analytical methods used during this intercomparison should be viewed as complementary, with each having its own merits (e.g. capacity for redox measurements (Bowie et al., 1998) and organic complexation determinations (Croot and Johansson, 2000)), although it is imperative that the availability of organically bound iron be ascertained as part of the routine determination of dissolved iron.

Improvements in our understanding of how biogeochemical processes mediate the iron available to each
analytical method will only occur through additional intercomparison exercises held both at sea and in shore-based laboratories. Previously, the lower reported trace metal values were believed to be the most reliable, since higher values were presumed to be due to contamination. This study highlights that this may not be the case. Investigators must demonstrate a willingness to examine the chemical intrusiveness of their rapid, real-time FI and CSV methodologies and determine which fractions of iron known to be present in seawater are analytically available. This task will require careful experimental design and an improved knowledge of, in particular, the extraction of organically bound iron during column preconcentration. Future FI methods may require on-line UV digestion in combination with sample acidification to determine the “total” dissolved iron fraction in the <0.2-μm size-fraction.

During a January 2000 workshop held to advance the certification of iron in seawater, SCOR-IUPAC Working Group 109 recommended that a first step towards a global intercomparison exercise would be the collection of a large volume sample, low in dissolved iron, which would be distributed worldwide to expert laboratories. This would eventually lead to the production of a CRM suitable for open-ocean iron measurements. During the intercomparison expedition reported here, with the ship located in a “low” iron region far from the coast, ~700 l of filtered surface seawater was collected in a cubic tank, acidified, mixed and subsampled into 1-l bottles. These samples were subsequently distributed to the 30 laboratories participating in the SCOR-IUPAC exercise for laboratory analysis. The stability and homogeneity of this set of intercalibration samples has yet to be determined and data from this exercise will be reported elsewhere in the future. Those wishing to obtain samples of this standard seawater material in order to aid method development, validation or to calibrate new instruments are invited to contact the corresponding author for further information.

Acknowledgements

We are grateful to the Alfred-Wegener-Institut für Polar-und Meeresforschung for providing a research platform on-board RV Polarstern for the intercomparison. Thanks to Loes Gerringa for help with the analysis of the TAC blank, Simon Ussher and Denis Mackey for internal review, and three anonymous reviewers whose comments improved the manuscript. This work was funded by the EU programme IRONAGES (Iron Resources and Oceanic Nutrients-Advancement of Global Environment Simulations, EVK2-1999-00031).

References


